Effects of Results-Based Financing on Patient Satisfaction in Afghanistan

By

M. Essa Tawfiq

A dissertation

submitted to the Victoria University of Wellington

in fulfilment of the requirements for the degree of

Doctor of Philosophy

30 September 2015

Table of contents

Abstract

A Results-Based Financing (RBF) program has been implemented in Afghanistan since September 2010 to improve the quality of health care and increase the utilization of maternal and child health services. This PhD study examines the impact of RBF on patient satisfaction and on determinants of patient satisfaction at health facility level in Afghanistan. Determinants of patient satisfaction in the study refer to health provider performance, availability of medicines, vaccines, equipment, and physical appearance of health facilities.

I used data collected from a panel of health facilities over a three-year period. The data consist of nearly 3000 patient observations and exit interviews. I included 112 health facilities in my study. These health facilities were part of the 428 health facilities which had been randomly assigned to treatment and control groups prior to the start of RBF in 2010. Financial incentives were distributed among health providers in the treatment facilities through four administration mechanisms: salary-based, task-based, equal-amount, and mixed-method. Follow-up surveys were conducted in 2011 and 2012 in the same 112 facilities, but for new cross-sections of patients and health providers. I analysed a range of patient satisfaction and patient satisfaction determinants measures using a regression-adjusted difference-in-differences estimation model.

The results from this study show that after a period of two years, there was an increase of only 8 percentage points in the proportion of patients who were very satisfied with services as a whole. However, the effect was not statistically significant. Similarly, specific aspects of patient satisfaction were not significantly affected by the intervention. Likewise, RBF did not have any significant effect on health provider performance, on availability of medicines, vaccines, and equipment, and on physical appearance of health facilities over a two-year period. I also found no difference in RBF treatment effects by the different incentive administration mechanisms.

My study provides evidence which suggests that paying monetary incentives alone may not have the impetus to improve health provider performance to the satisfaction of patients in a post conflict country. In such settings, RBF initiatives need to include both financial and non-financial incentives for health providers in order to achieve the intended objectives of quality of care and patient satisfaction. My study provides pragmatic recommendations aimed at holistic approaches to improving quality and delivery of healthcare in a post conflict setting.

Acknowledgements

I would like to thank the Victoria University of Wellington: The Victoria Doctoral Scholarship Program for financial support of my dissertation work. I am grateful to the School of Government of the Victoria University of Wellington for the substantial technical and logistics support I have received during my PhD work.

I would like to acknowledge the Afghanistan Ministry of Public Health (MOPH), and the Kabul Office of Johns Hopkins University and the Indian Institute for Health Management Research for sharing the databases and related research material that I have utilized in my dissertation. I would like to thank Dr.Najibullah Oshang, the MOPH RBF Focal Point who provided me with the database related to incentive administration, and who updated me on the program from time to time. I thank Dr. Ahmad Shah Salehi, Director of Healthcare Financing of the MOPH, who authorized me to use data from the databases available at the MOPH and at the Kabul Office of Johns Hopkins University in Afghanistan.

Many thanks are due to Dr. David Peters, Dr. Gilbert Burnham, Dr. Anbrasi Edward, Dr. Kayhan Natiq, Dr. Cyrus Engineer and many other colleagues from Johns Hopkins University who provided enormous technical and managerial support during the design, implementation and evaluation of the National Health Services Performance Assessment (NHSPA) and RBF programs in Afghanistan. I deeply appreciate the assistance of my seniors and colleagues from the Indian Institute for Health Management Research who continuously provided valuable insights into the design, implementation and evaluation of the NHSPA and RBF programs in Afghanistan. I am particularly grateful to Anubhav Agarwal, Rajpal Singh Yadov, and Suhail Ismail Shiekh with whom I was managing the day to day operation of the RBF and NHSPA programs for the two years I was working as a program manager with the Johns Hopkins University's Field Office in Kabul, Afghanistan.

Many thanks are needed for my Afghan colleagues in the Kabul Office of Johns Hopkins University for their professional support and friendship. They were working with me in the RBF and the NHSPA programs. Thanks are also due to the many Afghan men and women who participated in the NHSPA from which I used data in my dissertation.

I would like to express my heartfelt thanks and appreciations to my supervisors: Dr. Jaikishan Desai, and Dr. Dean Hyslop for providing me constructive feedback, clear advice, and continuous support. My supervisors have been the main driver in motivating me to carry on a *learning by doing* strategy, and to keep up my dissertation work on track and steady progress.

I am thankful to Professor Jackie Cumming, Professor David Hotchkiss, and Professor Steven Stillman for being the examiners, and for their time to review my dissertation and provide valuable insights and assistance to ensure my PhD work is academically sound and scientifically rigorous. I am grateful to Professor John Miller for chairing the oral defence of my dissertation where I received valuable advice and recommendations from the examiners to further improve the quality of my dissertation. I would like to extend my gratitude to Dr. William Newbrander – my former supervisor with the Management Sciences for Health (MSH) – who provided constructive comments on my research proposal in the first year of my PhD studies. Finally I would like to thank my wife and my five children for their loving support, and for encouraging me to complete my dissertation work within the allocated time.

Preface

Before joining Victoria University of Wellington for doing my PhD, I was working with the Johns Hopkins University in their Field Office in Afghanistan. I worked in the Results-Based Financing (RBF) and National Health Services Performance Assessment (NHSPA) programs for two years. I worked for nearly a year as the RBF program manager. I was involved in the establishment of a validation system based on which health facilities' reports were verified to receive the RBF incentives. I was also involved in the implementation of the RBF baseline household survey in late 2010. With respect to NHSPA, I managed the program for over a year, and I was involved in the revision of survey questionnaires, development of survey manuals, and various administrative tools. I was overseeing the implementation of NHSPA in 2011, and was involved in the day to day operation of the program which included sampling of facilities, deployment and monitoring of survey teams, data collection, and data analysis.

During the two years of my involvement in RBF and NHSPA, I became interested in studying the effects of RBF on patient satisfaction. With the permission and support of Dr. David Peters, Dr. Ahmad Shah Salehi and many other colleagues from the Johns Hopkins University, the Indian Institute for Health Management Research, and the Afghan Ministry of Public Health, I embarked upon PhD research to explore the impact of the RBF program.

Chapter 1: Introduction

This chapter first describes the organization of the dissertation. Then it provides background information on patient satisfaction and Results-Based Financing (RBF) programs in developing countries. The chapter then presents the Afghanistan healthcare system, and moves on to describing the RBF implementation in Afghanistan. In the later part, it discusses the research context, and lays out the study objective and research questions.

1.1 Organization of the dissertation

In this dissertation I am investigating the effects of a RBF intervention on patient satisfaction and on patient satisfaction determinants at health facility level in Afghanistan. Over two years, a RBF intervention has been implemented in health facilities where monetary incentives have been paid to health providers based on their performance in terms of utilization and quality of maternal and child health services. My motivation behind this PhD research is to see whether quality of care in terms of performance of health providers, availability of resources such as medicine, vaccines, equipment, and appearance of health facilities improve, and whether the level of patient satisfaction increase as a result of the intervention. I also compare the RBF treatment effects on the above outcomes between the four incentive administration types implemented in health facilities, and identify differences in effectiveness by the admin type.

The dissertation is organized into nine chapters:

Chapter one is mainly about identifying the research gap which becomes the focus of my dissertation. The chapter discusses background information about contracting approaches in the health sector in Afghanistan, and moves on to identify the research gap in the context of RBF in Afghanistan. Then the chapter describes the research objective and questions.

Chapter two reviews the relevant literature. The literature – reviewed here – is related to the effectiveness of RBF and similar programs in the health sector in developing countries, and to the research undertaken on patient satisfaction and quality of care. The chapter also discusses the literature related to health worker motivation.

Chapter three describes the four types of incentive administration, the target setting process, and the proportion of health workers' income from the program. The type of health services targeted and the amount allocated for target services as well as the incentive size and formula for computing incentives will be discussed.

Chapter four presents the conceptual framework upon which the study is based, describes the study design, sampling and data collection processes, including validity and reliability of the data collected. Chapter four also describes the estimation model, choice of regression models, outcome and control variables, and the measures undertaken to address the clustering effects.

Chapter five compares main characteristics related to health facilities, health workers, and patients between treatment and control facilities during the baseline period. The purpose is to check whether a similar distribution of covariates existed between the intervention and control groups prior to the start of RBF.

In chapters six, seven, and eight I estimate difference-in-differences model to measure the impact of the intervention. Chapter six discusses the RBF overall treatment effects on patient satisfaction. Chapter seven compares RBF effects on patient satisfaction disaggregated by the incentive administration types, and chapter eight investigates the RBF treatment effects on health provider performance, which is measured by patient – provider interaction, on availability of resources, [1] and on physical appearance of health facilities. The purpose in chapter eight is to study whether RBF has any impact on key structural and procedural factors which are essential for the delivery and quality of healthcare and they have positive impact on patient satisfaction.

Chapter nine summarizes the major findings based on the background information, data, and analyses presented in the dissertation. The chapter discusses the strengths and limitations of the study, describes the program and policy implications of the research findings, highlights the contributions of the study to the literature, and proposes areas for future research.

[1] In this dissertation, the terms 'health providers' and 'health workers' refer to all types of medical personnel (e.g. doctors, midwives, nurses, lab technicians, pharmacists, and vaccinators) working in health facilities.

1.2 Background

In recent years there has been increasing interest in evaluating patient satisfaction and quality of care in Low and Middle-Income Countries – LMICs (Bernhart et al., 1999; Roa et al., 2006; Andaleeb et al., 2007; Hansen et al., 2008; Aldana et al., 2011). The growing interest may be due to the growing pressure from the donor societies on the host nations to improve quality of healthcare and meet patients' legitimate expectations (De Geyndt, 1995; The World Bank, 2003). Poor quality of care has serious consequences for the health outcome (Andaleeb et al., 2007). The problem of low utilization of key health services is strongly linked to poor quality of care in LMICs (Reerink & Sauerborn, 1996). Addressing patients' legitimate expectations in its own right is considered to be an ultimate goal for any healthcare system (Evans et al., 2001).

Performance-based contracting of health services and RBF is part of a global effort to improve quality of care and increase utilization of health services in LMICs (Loevinsohn & Harding, 2005; Meessen et al., 2006; Liu et al., 2008; Oxman & Fretheim, 2009). Influenced by the donor societies, health authorities in several LMICs particularly the ones in post conflict and fragile states due to the low capacity of public health sector adapted contracting approaches of health services to expand the delivery of mainly primary healthcare to their citizens (Loevinsohn & Harding, 2005; Palmer et al., 2006; Batley & Mcloughlin, 2010). Some other LMICs were intrigued by the encouraging results obtained from various models of pay for performance (P4P) implemented in the industrialized nations, and under such influences they started to replicate results-based financing approaches in their countries (Eldridge & Palmer, 2009; Ridde, 2005).

Results-Based Financing (RBF) is defined as the "transfer of money or material goods conditional on taking a measurable action or achieving a predetermined performance target" (Eichler, 2006; Oxman & Fretheim, 2009; Witter et al., 2012). Other authors use a different term for defining the same concept. For example, according to Meessen et al. (2011), performance-based financing (PBF) is a mechanism by which health providers are, at least partially, funded on the basis of their performance. The terms RBF, pay-for-performance (P4P), and PBF have been defined in the same way in the literature. For example, Eichler (2006); Lindahl et al. (2008); and Witter et al. (2012) use the same definition (above) for PBF, P4P, and RBF. Lindahl et al. (2008) state that the terms RBF and P4P are often used

interchangeably. In a recent paper, Gorter et al. (2013) state that RBF refers to all types of mechanism which link payments to results and performance. The last definition mainly focuses on the supply side where provider payments either partially or totally are conditional on their performance.

According to Eichler (2006), there are three types of RBF: supply side financing, demand side financing, and a mixed type of the two. The supply side interventions are designed to change behaviour of health providers, while the demand side intends to address barriers that prevent households and individuals from using public health services. The hybrid form of the approach aims at addressing problems on the demand side and supply side simoutenously. The focus of RBF in Afghanistan is on the supply side to provide financial incentives for health providers to improve performance and quality of care in order to increase utilization of maternal and child health services in public health facilities.

1.3 Afghanistan healthcare system

In Afghanistan, structure of the healthcare system has a lot to do with contracting and performance-based programs in the recent one and half decades. This is due to the fact that before the fall of the Taliban, no coherent healthcare system existed and health services were provided by a multitude of national and international non-governmental organizations (NGOs) in a fragmented manner (Bower, 2002; Sondrop, 2004; Bristol, 2005; Strong et al., 2005). The public healthcare system was largely dysfunctional, and the NGOs often focused on their own areas of interest. Health facilities, health workers and other healthcare resources were insufficient in quantity and quality in the country. Inheriting a dysfunctional system from the Taliban's period, the public health sector did not have the capacity to provide health services in an equitable and efficient manner in the country (Newbrander et al., 2014).

After the fall of the Taliban and establishment of an interim government in December 2001 which was followed by the influx of international aids, the Afghan government put efforts to rapidly expand coverage of basic health services in the country (Strong et al., 2005; Palmer et al., 2006). In early 2002, the Afghan Ministry of Public Heath (MOPH) stated its intent to address the fragmentation in the health sector by implementing a common package of health services in the country (Strong et al., 2005; Newbrander et al., 2007). The decision was taken following the joint donor mission which had the intention to develop the Basic Package of

Health Services – the BPHS (Afghanistan Ministry of Public Health, ,March 2003), and to expand the coverage of BPHS through providing grants and contracts to NGOs (Strong et al., 2005; Waldman et al., 2006; Arur et al., 2010).

This was a historical change in the rebuilding of the Afghanistan healthcare system, as the country began to experience the purchaser-provider split in the health system where the MOPH would contract NGOs to deliver health services to Afghan citizens on behave of the MOPH. The purchaser-provider split is a service delivery model in which third-party payers are kept separate from service providers. The operation of providers is managed by contracts. One of the main aims of purchaser-provider split is to create competition between providers. Competition and other incentive structures built into the contractual relationship are believed to lead to improvement in service delivery, such as cost effectiveness, efficiency, and improved quality to the satisfaction of community members (Tynkkynen et al., 2013).

The change raised concerns among government authorities as they felt that healthcare delivery was the state's function (Newbrander et al., 2014). It also raised concerns among the NGOs as they perceived they would lose their independence because of government involvement. With the donor support, the MOPH started to take a stewardship role of which building a public-private partnership and contracting of health services was a major part of it (Newbrander et al., 2014). In March 2003, the MOPH and its technical partners finalized the BPHS, upon which the MOPH was able to contract with NGOs and legally bind them to the provision of a predefined set of basic health services (Newbrander et al., 2007) . The development and implementation of the BPHS was one of the major factors contributing to the rebuilding of healthcare system in Afghanistan (Newbrander et al. 2014; Dalil et al., 2014). In 2005, the MOPH complemented the BPHS by finalizing the Essential Package for Hospital Services – the EPHS (Afghanistan Ministry of Public Health, ,2005). The EPHS further strengthened the healthcare system, as the MOPH contracted with NGOs to provide secondary and tertiary healthcare besides the delivery of primary healthcare as specified in the BPHS (Newbrander et al. 2014). Both packages also specified the number and type of health facilities, health personnel, equipment and medical supplies, and essential drugs required for the healthcare delivery. Moreover, both packages described the organization, and relationships between various structures involved in the Afghanistan healthcare system.

In 2003, the MOPH with support from the World Bank established a new unit within the Ministry to initially manage contracting with NGOs for the BPHS delivery through the PPA (Performance-based Partnership Agreements) contracting approach.[2] This new unit was called the Grant and Contract Management Unit (GCMU) which was instrumental in the expansion of the BPHS in Afghanistan (Waldman et al. 2006). Throughout the first few years, the GCMU's capacity was built, and it started to manage grants and contracts funded by other donors such as the USAID and European Commission in Afghanistan (Waldman et al. 2006). The GCMU is currently managing almost all grants and contracts pertaining to health service delivery in Afghanistan (Dalil et al., 2014).

The MOPH used both contracting-in and contracting-out approaches to expand the delivery of BPHS and EPHS services (Waldman et al., 2006). Contracting-in refers to the contracting model where the MOPH contracts with health workers in health facilities. The MOPH has overall responsibility for managing healthcare provision in the geographic area under the contract. Contracting-out refers to the contracting arrangement where the MOPH contracts with NGOs who in turn contract with health workers in health facilities. The NGOs have overall responsibility for managing the delivery of healthcare to community members living in the geographic areas covered by the contract. In 31 provinces, the NGOs were contracted-out, and in the 3 remaining provinces the MOPH Strengthening Mechanism (SM) was contracted-in to provide health services (Palmer et al., 2006). SM is the contracting-in model of service delivery, and it has been active in three provinces in north of Kabul province since 2004. The purpose of establishing the SM was to compare this model of service delivery with the contracting-out model. Several studies have compared the utilization and quality of health services between the contracting-in and contracting-out approaches in Afghanistan (Hansen et al., 2008; Arur et al., 2010).

The NGOs and SM were provided technical assistance to improve their capacity for grant and contract management by the GCMU and by the Management Sciences for Health (MSH).[3] Despite serious concerns raised by some scholars such as Ridde (2005) over the government capacity, the GCMU demonstrated that it has the capacity to manage contracts and grants in

[2] PPA has been financially supported by the World Bank in over one third of Afghanistan. Provision of the BPHS and EPHS services in the other two thirds of Afghanistan has been financially supported by the USAID, and European Commission.
[3] MSH is an international organization which has been involved mainly in grant management of USAID funding, and capacity building programs in the health sector in Afghanistan since 2002.

the health sector in Afghanistan (Arur et al., 2010; Dalil et al., 2014; Newbrander et al., 2014), and it seems likely that contracting approaches in the health sector continue in Afghanistan at least in the medium term (Siddiqi et al., 2006; Sabri et al., 2007; Dalil et al., 2014).

Substantial investments in the health sector have led to remarkable improvements in maternal and child health status in recent years in Afghanistan. For example, the 2010 Afghanistan Mortality Survey estimates that there are 327 (260 – 394)[4] maternal deaths for every 100,000 live births and 97 death before the age of five years for every 1000 children born (Afghan Public Health Institute (APHI) of the Ministry of Public Health, 2010,). According to a sub-national survey in 2003; however, the maternal morality ratio (per 100,000 live births) was estimated 418 (235-602) in Kabul, 774 (433 – 1115) in Laghman, 2182 (1451 – 2913) in Kandahar, and 6507 (5026 – 7988) in Badakhshan (Bartlett et al., 2005). According to Bartlett et al. (2005), the death rate in children under 5 years per 1000 live births was 109 (87 – 130) in Kabul, 190 (149 – 230) in Laghman, and 323 (266 – 378) in Badakhshan. The significant reduction in the mortality rates is consistent with improvement in various determinants of health, including an increasing age at marriage, higher use of contraceptives, lower fertility rate, better immunization coverage, improvement in safe motherhood services (such as antenatal care, deliveries, and postnatal care), involvement of community health workers and increasing access to the BPHS and hospital services (Rasooly et al., 2014).

While some scholars such as Newbrander et al. (2014); Dalil et al. (2014) assert that achievements in rebuilding the healthcare system in Afghanistan is exemplary, other scholars such as Michael et al. (2013) show serious concerns about the credibility of the achievements attributable to contracting approaches in Afghanistan. According to Newbrander et al. (2014); Dalil et al. (2014), the achievements are more prominent in terms of building the stewardship role of the MOPH and establishing public – private partnership approaches. The authors continue their statements by highlighting the capacity building approaches of Afghan NGOs who substantially have contributed to the expansion of healthcare coverage in the country. On the other hand, Michael et al. (2013) show doubt about the accuracy of these achievements by questioning the feasibility of public–private partnership at a large scale in a fragile state such as in Afghanistan. Michael et al. (2013) continue their contentions by pointing to the

[4] The values inside the brackets refer to the 95% confidence interval.

perceived lack of transparency in the process of contracting, and sustainability of the approach. The authors further argue that the notion that implementation of BPHS and EPHS would contribute to state-building has not happened, and that while health services through public facilities is largely accessible throughout the country, it compares unfavourably with the private sector in terms of quality, and this has restricted the utilization of health services.[5]

Some other critics of contracting approaches such as Ridde (2005) question the superiority of private over public service provision. The author focuses on analysing contracting-out approach, and raises serious concerns over the privatization of healthcare delivery in Afghanistan, and proposes ways for the international donor community to strengthen the Afghan government to get involved in the direct healthcare provision in order to build its statehood and stability in the country.

Considering the escalation of insurgency in recent years (Giustozzi, 2009; Minami, 2010), the approach suggested by Ridde (2005) may not seem to be an appealing undertaken. In recent years, large swathes of the country have been engulfed in an open conflict and are largely controlled by the insurgency (Giustozzi, 2009; Minami, 2010).Therefore, except for the for-profit private sector and NGOs, no government entities are tolerated by the insurgency for the provision of public services such as healthcare and education. The ongoing violence and conflict has been cited as the top factor limiting access to healthcare services in Afghanistan (Carthaigh et al., 2014). The authors raise serious concerns over the partiality of healthcare system because of linking the delivery of health services to political goals such as state-building and stabilization in Afghanistan. Carthaigh et al. (2014) argue that allocation of health resources should be needs-based, not politicized, because impartial healthcare delivery is vital to increasing access to basic and life-saving care in the country.

Afghanistan is one of the poorest countries in a post conflict, and yet fragile state where effects of contracting of health services and performance bonus have been identified as significant in terms of equity in access, and utilization of health services (Loevinsohn & Harding, 2005; Strong et al., 2005; Arur et al., 2010; Dalil et al., 2014; Waldman & Newbrander, 2014). In addition to the contract amount, performance bonus has been paid to

[5] State-building refers to the creation of new governmental institutions and strengthening of existing ones which have the potential to influence community members' perception that the Afghan government has strengthened its capacity in terms of providing public services and enforcing rule of law in the country.

the best performing NGOs or SM in Afghanistan (Hansen et al., 2008). According to Hansen et al. (2008), the NGO or SM was paid a performance bonus of 1% of the contract amount if they achieved an increase of 10 points in the mean score across the 29 indicators on the Balanced Scorecard (BSC).[6] The BSC was used as a device for setting performance benchmarking on various domains of healthcare system amongst the 34 provinces of the country (Peters et al., 2007).

1.4 Results-based financing in Afghanistan

Having the experience of employing the PPA performance bonus, the MOPH decided to launch the RBF intervention where health facilities became the target for receiving the performance incentives. The RBF initiative to pay incentives to health facilities was an attempt to address the low motivation of health professionals which was thought to be a major cause of poor quality and low utilization of maternal and child health services. The main difference between performance bonus and RBF (with the current design) is that under the former (i) health facilities assessment was done annually, (ii) performance bonus was paid only to the best performers, (iii) performance was assessed on various aspects of healthcare delivery and quality, and (iv) performance was evaluated at the province level. Unlike the PPA performance bonus, under the RBF program, (i) performance was assessed at health facility level, (ii) incentives were paid based on the extra number of target health services,[7] (iii) performance incentives were paid to health facilities every three months, and (iv) no additional managerial or logistic support was provided to health facilities.

In Afghanistan RBF started in September 2010 in 9 provinces and by December 2010, it was extended to 11 provinces out of the 34 provinces of the country. The project implementation period was divided into three years: September 2010 to August 2011; September 2011 to August 2012; and September 2012 to March 2013 (Figure 1.1).

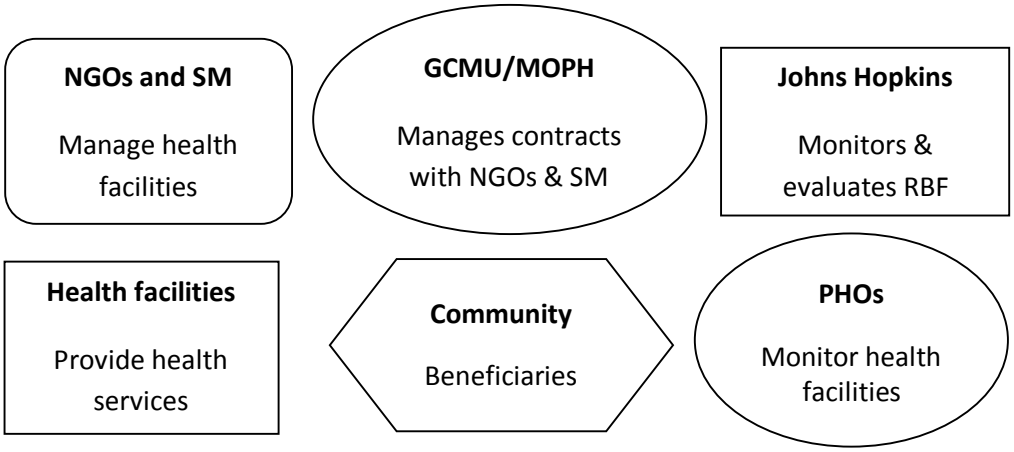
[6] The BSC is a health performance assessment tool, produced by the third party (Johns Hopkins University's Kabul Office) and used by the MOPH.
[7] Some adjustments have been made to account for the overall quality of care which was assessed by the use of "national monitoring checklist" in health facilities.

Figure 1.1 Start dates of RBF implementation in the eleven provinces

Provinces		Number of facilities	First Year			Second Year		Third Year	
Names	Number		Sept 2010	Dec 2010	Aug 2011	Sep 2011	Aug 2012	Sep 2012	Mar 2013
Panjshir, Samangan, Parwan, Bamyan, Saripul, Kunduz, Kandahar, Balkh, Jawzjan	9	174							
Takhar, Badakhshan	2	56							
Total	**11**	**230**							

In Takhar and Badakshan provinces due to delays in the contract modification, the RBF program started in December 2010

There were several key stakeholders involved in the design, implementation, and monitoring and evaluation of RBF in Afghanistan. The funding, as part of a global effort, was provided by the Norwegian government via the World Bank. On the ground, the MOPH plays the role of policy maker and regulator, GCMU/MOPH manages the contracts, The Johns Hopkins University monitors and evaluates the program, NGOs and SM manage health facilities, health facilities provide health services, and Provincial Health Officers (PHOs) perform the leadership role of MOPH at provincial level. The RBF stakeholders' roles are displayed in Figure 1.2 below.



To launch RBF, the MOPH amended the existing contracts of NGOs and SM for providing health services in accordance with the BPHS guideline in the 11 provinces. According to the specification of the existing contracts, health facilities were responsible for offering the BPHS services, which included the maternal and child health services targeted by the RBF program. The initiation of RBF was in response to the low coverage of important preventive, promotional, and curative services which remained low by global standards, in spite of considerable gains in the coverage, utilization, and quality of health services since 2003 in Afghanistan. For example, in rural Afghanistan the use of modern contraceptives was 16%, the use of antenatal care was 32%, the proportion of women in labour who were attended by skilled birth attendants was about 19% in 2006 (Afghanistan. Ministry of Public Health, 2008). Diphtheria Pertussis Tetanus (DPT) vaccination coverage, which is commonly used as a measure of the effectiveness of the routine healthcare system in delivering immunization services, was over 60% for DPT1 among children 12-23 months old. However, with the second dose of DPT there was a 12 percentage points drop in coverage, and with the third dose there was a further 14 percentage points drop to 34%. These declines in DPT coverage indicated that there were opportunities missed by the healthcare system in Afghanistan.

With the objective of increasing utilization and quality of maternal and child health services, RBF monetary incentives paid to health facilities were distributed among health workers according to one of the four administration types: salary-based, task-based, equal-amount, and mixed-method. Under the salary-based incentive administration, each staff member (technical and support) has received incentives proportional to the share of his/her monthly salary in total salary payment. Under the task-based incentive administration, the health worker who was involved in the provision of the RBF target services received the incentives based on his/her direct involvement in the delivery of the target services. Under the equal-amount incentive administration, the incentives were distributed equally among the staff member, and under the mixed-method administration, both salary-based and task-based administration was practiced in a health facility (These are discussed in chapter 3).

1.5 Research context

There is growing concern that due to high maternal and child mortality some developing countries may not be able to achieve the health-related Millennium Development Goals (MDGs) particularly MDG4 and MDG5 (Bhutta et al., 2004; Hogan et al., 2010; Travis et al., 2004). MDG4 is related to the reduction of child mortality, and MDG5 is about the improvement of maternal health in LMIC settings. Despite its recent achievements in the reduction of maternal and child mortality, Afghanistan has not reached the targets of MDG4 and MDG5 yet (Islamic Republic of Afghanistan). To achieve these targets in developing countries, one of the approaches being tried is RBF or performance-based financing. According to Soeters et al. (2006), RBF is a special type of contracting health services. RBF and/or contracting of health services has the potential to expand coverage, increase utilization, and improve quality of healthcare in developing countries (Bhushan et al., 2002; La Forgia, 2005; Soeung et al., 2006; Eichler et al., 2007; Basinga et al., 2011).

According to Oxman and Fretheim (2009) RBF is a simple concept but, its actual implementation can vary according to (a) the level at which the incentives are targeted (e.g., individual health providers, health facilities, health sector organizations); (b) the targeted results (e.g., utilization of services, quality of care, equity in access to care); (c) the performance indicators (e.g., what is measured, how it is measured, who measures it); (d) the performance targets (e.g., pay per unit of result or pay only if a target is achieved); (e) the type and size of incentives (e.g., cash, vouchers or material goods with the frequency of transfers); (f) the proportion of financing that is paid based on achieving results; and (g) the supplementary support provided to facilities (e.g., availability of more resources, training, monitoring, feedback, regular supplies of essential drugs, salary increment, building new facilities, better governance and decentralization, and involving stakeholders).

Implementation of RBF in the health sector has attracted keen attention of health researchers, managers, policy makers, and donors worldwide. Currently there are serious debates among scholars on the effectiveness of RBF in the health sector in LMIC countries. For example, the assumed positive impact of RBF on patient satisfaction suggested by Meessen et al. (2011) and the possible unintended consequences of RBF on patient satisfaction raised by Ireland et al. (2011) have remained unexplored. According to Meessen et al. (2011), since there is a strong link between performance of health workers and the amount they receive, it is likely

that RBF can increase the quality of care leading to higher levels of patient satisfaction. In contrast, Ireland et al. (2011) argue that RBF can turn the attention of health providers towards clinical quality of care at the expense of patient values and expectations, resulting in lower levels of patient satisfaction.

Patient satisfaction is a subjective measure, and is different from patients' reports about objective attributes of care. Patients, for example, might be asked to report whether they were told about possible side effects of drugs. This gauges patients' objective assessment of the quality of care. Alternatively in a subjective way, patients might be asked "how satisfied they were with the way health providers explained the possible side effects of the drugs prescribed". The subjective evaluation of specific domain of care refers to patient satisfaction with the domain in question. In this study I study the impact of RBF on overall patient satisfaction and on patient satisfaction with various aspects of healthcare delivery.

Patient satisfaction is very important as research shows that consequence of patient dissatisfaction can be grave on the health outcome. It has been found that patient satisfaction can affect choice of provider, utilization of care, and compliance with treatment (Andaleeb, 2001; Meessen et al., 2011). In a study in Afghanistan, Ameli and Newbrander (2008) identified a significant positive association between patient satisfaction and the use of curative health services. Similarly, the consequences of patient dissatisfaction can be serious on the health outcome. Patients who are dissatisfied with health services may not adhere to the treatment and to the follow-up visits, and may share their negative experience with others and dissuade them from using a particular source of healthcare (Andaleeb et al., 2007).

Another crucial consequence of patient dissatisfaction in some LMIC settings concerns the relatively large amounts of out-of-pocket expenditures incurred by patients to obtained health care in neighbouring countries (Andaleeb, 2001; Ashrafun & Uddin, 2011). For example, Ashrafun & Uddin (2011) assert that due to poor quality of care in hospitals, Bangladeshi patients have to travel to the neighbouring countries to get quality hospital healthcare. A similar concern was expressed by the MOPH that Afghan patients spend huge amount of money to get better quality services in neighbouring countries as they are not satisfied with the quality of hospital services they receive inside Afghanistan.

1.6 Research objective and questions

The objective of the study is to investigate the effects of RBF intervention on patient satisfaction in Afghanistan. The objective stems from the assumption that RBF motivates health providers to improve performance and provide quality services to patients (Henderson & Tulloch, 2008; Basinga et al., 2011; Meessen et al., 2011), and this translates to higher levels of patient satisfaction with the health services offered by health facilities. Based on this assumption the following three questions will be addressed.

1. What is the RBF treatment effect on patient satisfaction?
2. What is the RBF treatment effect on health provider performance?
3. What is the RBF treatment effect on availability of resources,[8] and on physical appearance of health facilities?
4. Is there any difference in the RBF treatment effect on patient satisfaction, health provider performance, and on availability of resources, and on physical appearance of health facilities under the four incentive administration types?

The first three questions attempt to examine treatment effects on patient satisfaction and on patient satisfaction determinants. The last question is an exploratory one which explores between the admin types with the objective of finding any difference in the treatment effects by the admin types, as research shows that provider payment mechanism can affect their performance (Robyn et al., 2012). The first question includes both overall satisfaction which refers to general assessment of various aspects of care by a patient, and questions on specific aspect of care. Knowing treatment effects on patient satisfaction with aspects of care is very important, as according to Rao et al. (2006) for quality improvement interventions, it is paramount to obtain actionable information on patient satisfaction with specific aspect of care rather than on overall satisfaction with care alone. The second and third questions relate to those aspects of quality of care that are potentially influential to patient satisfaction.

[8] Availability of resources in this study refers to the availability of medicine, vaccines, and equipment in health facilities

Chapter 2: Literature review

At first, this chapter reviews and discusses the literature related to RBF and contracting of health services in LMICs. Then the studies on patient satisfaction in LMICs will be reviewed and discussed. The purposes are: (i) to provide an overall picture of the effectiveness of performance-based and RBF approaches in the health sector in developing countries, and (ii) to identify potential factors that influence patient satisfaction in the context of LMICs.

2.1 Results-based financing and contracting

There is growing interest in studying the effects of RBF and contracting on various aspects of healthcare. In LMIC countries, the intervention's effects have largely been studied in relation to utilization of, access to, quality of, and patient satisfaction with healthcare. These studies have been conducted in a wide range of contexts and have had different evaluation designs. Overall the findings from these studies provide evidence which indicates that RBF and similar approaches, if designed and implemented properly, can achieve at least short-term results (Soeters & Griffiths, 2003); Loevinsohn & Harding, 2005; Strong et al., 2005; Meessen et al., 2006; Palmer et al., 2006; Meessen et al., 2007; Basinga et al., 2011).

The important role of content and context in relation to the success of RBF and similar interventions has been emphasized in the health sector. In the literature, the aspects relating to content such as the design and specification of contracts, performance targets, selection of indicators, levels at which incentives are targeted, size of incentives, and supplementary elements are considered very important in the success of a health intervention (Liu et al., 2008; Oxman & Fretheim, 2009). Similarly, contextual factors such as management capacity and political will of host government, commitment of donor society, security situation, and managerial capacity of implementing organizations (e.g. NGOs) are considered crucial in the success of RBF and contracting approaches (Liu et al., 2008; Oxman & Fretheim, 2009; Morgan, 2010).

A large body of literature related to RBF and contracting in LMICs focuses on the utilization of and access to primary health care. The findings from these studies are largely but not consistently in favour of RBF and contracting in the health sector. A great number of studies have found substantial increase in the coverage of maternal and child health services (Soeters & Griffiths, 2003; Danel & La Forgia, 2005; Sekabarga et al., 2011; Olken et al., 2012).

Soeters & Griffiths (2003) used a before-and-after design with no control group, and investigated the effects of a contracting-in approach on the utilization of maternal and child health services in a district in Cambodia. Using data collected from household surveys, the authors reported significant increases in the use of institutional deliveries,[9] antenatal care, family planning, immunization, and use of ORS for diarrhoea in the intervention district.

In the study by Danel & La Forgia (2005), the authors used a before-and-after design and compared the utilization of maternal and child health (MCH) services between three delivery models – direct model, mixed model, and traditional model – in the rural areas of Guatemala. Unlike the traditional model, under the direct and mixed models, NGOs were contractually responsible for providing health services and acting as financial managers on behalf of the Ministry of Health. Using data from household surveys, the authors found considerable increase in the coverage of immunization, utilization and quality of postnatal care, and use of ORS for the episodes of diarrhoea in the areas covered under the direct and mixed models.

In Rwanda, Sekabarga et al. (2011) used a longitudinal design, and studied the effects of PBF on the utilization of MCH services over five and seven years. Using data from nationwide population-based survey, the authors reported significant increases in the utilization of maternal and child health services, particularly the assisted deliveries among the poorest from 12% in 2000 to 27% in 2005 and to 43% in 2007.

In Indonesia, Olken et al. (2012) used a randomized controlled trial of 3000 villages to investigate the efficacy of performance incentives in terms of improving 12 maternal and child health and education indicators. The authors found that over the two year project period, utilization of 8 targeted maternal and child health services was higher in the incentivised areas than in the non-incentivized areas.

Unlike the above, some other studies reported no effect of PBF on the use of MCH services (Morgan, 2010). In the study by Morgan (2010), it was hypothesized that performance-based contracting could improve health provider performance and increase utilization of MCH services in Uganda. The author used a quasi-experimental design, and used data collected through household surveys and routine reports of health facilities. After two and half years

[9] Institutional delivery is defined as delivery in a hospital, clinic, or health centre (Saxena et al., 2013).

and three survey rounds, it was identified that there was no increase in the utilization of MCH services in the facilities which received performance incentives. The author explored the possible reasons for the apparent failure of the program, and suggested ways that highlight the importance of design and implementation of PBF interventions in the health sector.

Several studies have reported significant increase in the use of curative health services (Bhushan et al., 2007; Arur et al., 2010; Soeters et al., 2011). Bhushan et al. (2007) used a before-and-after design with control group for five districts in Cambodia, and studied the effect of contracting-out on the coverage and utilization of health services, and on out of pocket health expenditure. The authors found significant improvement in the targeted outcomes in the intervention districts in comparison to the control districts between 1997 and 2003. Public health funding increased, and out of pocket health expenditure reduced as people switched from using traditional health providers to government health facilities.

In Afghanistan Arur et al. (2010) used a before-and-after design with control groups, and used data from household and facilities surveys. The authors reported significant increases in the use of outpatient visits in contracted facilities versus non-contracted facilities between 2004 and 2005. The authors found a significant increase of 29% in outpatient visits, and reported that the 29% increase in outpatient visits was disaggregated by an increase of 41% by female patients, an increase of 68% by the poorest quintile, and an increase of 27% by children aged under 5 years old over one year time. This implies that the program not only addressed the socioeconomic equity in healthcare delivery, but also affected the equity in terms of gender and priority age-group for basic and curative services in Afghanistan.

In the Democratic Republic of Congo, Soeters et al. (2011) investigated the effect of PBF on the quality and quantity of MCH services. Using a before-and-after design with control groups with data collected through household surveys, the authors found significant increase in the use of childbirth services, and moderate increases in the use of antenatal care, and immunization in the intervention facilities between 2005 and 2008. Also patient perceived quality of care significantly increased in the intervention facilities over the project period.

In other countries, significant increases in the use of MCH services as a result of contracting approaches were reported (Basinga et al., 2011; Eichler et al., 2007; Meessen et al., 2006; Rusa et al., 2009; Soeters et al., 2006; Soeung et al., 2006).

Soeung et al. (2006) used a before-and-after design of ten villages with no control groups in six provinces in Cambodia. These villages were selected by the health authorities because coverage of immunization was very low as the villages located in remote and hard-to-reach areas. A performance-based type of contracts were signed between the staff involved in the delivery of immunization services and the local health authorities. The authors used the routine coverage reports from the immunization centres, and reported significant increase in the vaccination coverage among children between 2002 and 2003.

In Rwanda, Meessen et al. (2006) used a before-and-after design with control groups, and studied the effectiveness of PBF on the use of MCH services. The authors identified considerable increase in the use of targeted services over a period of three years. In another study in Rwanda, Basinga et al. (2011) studied whether PBF can positively affect the utilization and quality of MCH services. The authors used randomized controlled trial of health facilities, and used data from household and facility surveys. After a period of 23 months, the difference between the baseline and end-line results showed significant increase in the use of institutional deliveries, and significant improvement in the quality of prenatal care (compliance with clinical guideline on prenatal care).

In the study by Rusa et al. (2009), the authors used a quasi-experimental design of health facilities, and used data from supervision reports of health facilities in Rwanda. The authors studied the effect of PBF on quality and quantity of MCH services over a three-year period, and found that utilization of health services considerably increased for the services that were less well-organized (these services were institutional deliveries and growth monitoring). Utilization of vaccination, and antenatal care was not affected positively; but utilization of family planning and curative services increased in both groups of facilities over time. Quality of care, defined as compliance with national and international norms, also significantly improved for all services in both groups of health facilities. In another study in Rwanda, Soeters et al. (2006) used a before-and-after design with no control group, and examined the effects of a PBF program on health facilities performance and out of pocket health expenditure. Using data from household surveys, the authors reported significant increases in the use of institutional deliveries and family planning services over the three year period.

In Haiti, Eichler et al. (2007) used a before-and-after design, and investigated the effect of a performance-based contracting approach on the utilization of MCH services. The program

initially (when started in 1995) was not designed as performance-based; however, in 1999 the donor decided to change it based on partial attaining performance targets that the existing NGOs had to achieve. Technical assistance alongside some cross fertilization activities between NGOs (e.g., networking and meetings) was provided. The authors reported significant improvement in the use of institutional deliveries, and immunization coverage over six years since the start of performance-based initiative. There were, however, moderate increase in the use of antenatal and postnatal care in health facilities over the project period.

Out of pocket health expenditure, which is considered as one of the major barriers preventing the poor from using healthcare services, has also received considerable attention. Several studies have found the evidence indicating significant reductions in the out-of-pocket health expenditure after the introduction of RBF and contracting in the resource scarce settings (Soeters et al., 2006; Soeters et al., 2011). In the study by Soeters et al. (2006), in Rwanda it was found that catastrophic health expenditure has significantly reduced as a result of PBF. The authors reported that out of pocket health expenditure decreased by 62% from US$ 9.05 to US$ 3.45 between January 2003 and October 2005. Within the same time period, illnesses with catastrophic expenditure dropped by 72% - showing a significant improvement.

In another study in the Democratic Republic of Congo, Soeters et al. (2011) found that health expenditure substantially decreased among the poor while it increased among the relatively wealthy group of the participating households. For example, the authors reported that per capita out of pocket health spending increased by 45% between 2005 and 2008; however, within the same time period the per capita out of pocket health spending decreased by 14% among the poorest 25 percent of the households.

Efficiency of contracting and productivity of health workers have been covered by a number of studies (Cercone et al., 2005; Meessen et al., 2007; Loevinsohn et al., 2009; Soeters et al., 2011). Cercone et al. (2005) used a before-and-after design comparing the effect of a contracting approach for general practitioners with traditional publicly-managed clinics in Costa Rica. The authors identified that contracted providers conducted significantly more visits per capita than the traditional clinics, and reported significant (30%) lower expenditure per capita in the contracted providers than in the traditional clinics.

In a PBF program in Rwanda, Meessen et al. (2007) investigated health workers productively. The authors used a before-and-after design with no control groups, and used data from health facility surveys. Meessen et al. (2007) reported higher productivity among health workers in the intervention facilities between 2001 and 2003.

In a study by Loevinsohn et al. (2009) in a district in Pakistan, the authors investigated the effects of a contracting-out program on the quality, quantity, and efficiency. The authors used a quasi-experimental design, and used data from household and facility surveys, and reported considerable improvement in the efficient use of resources by the contracted NGO which achieved important goals at the same cost to the government facilities. However, it was reported that the program had little effect on the coverage of preventive services, and on the clinical quality of care. Similarly, in the study by Soeters et al. (2011), it was reported that PBF substantially contributed to the generation of revenue, availability of qualified staff, and regularity of supervision in health facilities.

Effectiveness of RBF and contracting on the clinical quality of care has also received sufficient attention. In this respect, several studies have found significant improvement (Hansen et al., 2008; Huntington et al., 2010; Peabody et al., 2011; Rusa et al., 2009), while some other studies have reported moderate improvement (Basinga et al., 2011), and yet some others have found no evidence to suggest improvement in the clinical quality of care (Loevinsohn et al., 2009).

In a PBF program in 30 hospitals in the Philippine, Peabody et al. (2011) used randomization of health providers for receiving performance bonus. The clinical knowledge of health providers was assessed using a vignette approach of questions applied after 12 months, 18 months, 24 months, 30 months, and 36 months since the start of the intervention. The authors reported significant improvement in the clinical performance of health providers which can positively affect technical quality of care. In the study by Rusa et al. (2009) in Rwanda it was reported that compliance with national and international norms significantly increased as a result of the PBF program in the intervention facilities.

Success or failure of a PBF program mainly depends on the design choice, and the characteristics of the context in which it is implemented. The importance of design choice, implementation approach, and contextual factors has been emphasized in the literature. In a

literature review of 128 evaluation studies on the effects of P4P in the industrialized countries, Van Herck et al. (2010) concluded that P4P effects can be judged on a whole range of outcomes from being encouraging to disappointing. The authors asserted that the evidence in their review has provided further indications on how effect findings are likely to be affected by the design choice and context where the program is introduced.

Similarly the importance of design and implementation issues of P4P programs has been highlighted by Young et al. (2007) who discussed these issues by considering different approaches in light of several P4P programs in the United States. According to the authors, these issues include strategies for choosing targets, performance indicators, unit of accountability, incentive size, pay out formula, and collaboration among payers.

Paying close attention to these issues is also crucial for designing and implementing performance-based interventions in LMIC countries. In this regard, the literature review by Oxman & Fretheim (2009) is relevant. In a systematic review of several selected RBF and contracting programs in LMICs, and interviews with key informants, Oxman & Fretheim (2009) came up with the conclusion that when RBF schemes are used, they should be designed carefully, including the level at which incentives are targeted, the choice of targets and indicators, the type and size of incentives, the proportion of incentives that is paid for achieving the targets, and the ancillary support (e.g., supervision, training, essential drugs, equipment, etc.). Also type and complexity of health services targeted can have significant impact on the success or failure of a RBF program (Town et al., 2004).

Targeting the right level of healthcare delivery for receiving the incentives is crucial for the success of a RBF program (Oxman & Fretheim, 2009; Young et al., 2007). This is important because it can identify who is accountable for achieving the targets, and whether the intervention can achieve the intended outcome. In the UK, for example, most quality-based payments are made to the medical practices, not to individual providers (Smith & York, 2004). In the US, similarly, many programs pay the performance incentives to medical groups or physician organizations (Young et al., 2007). It is argued that paying performance incentives to medical groups or organizations can encourage team work and improve quality of care. On the other hand, some other scholars argue that paying physician organizations can minimize the power of incentives as the money can be distributed among the individual providers in ways they may not like (Bokhour et al., 2006).

A similar concern can be raised for the PPA contracts in Afghanistan where performance bonuses were rewarded to NGOs, not to health facilities, and it has remained unclear whether and how much of the bonus was distributed among health workers in health facilities. The above caveat might be one of the rationales to direct the performance incentives at health facility level in the current RBF program. However, whether targeting incentives at health facility level can improve the health outcome, and whether the different types of incentive administration employed in health facilities can influence the health outcome remain to be explored in this study.

One of the critical elements in the design of a RBF intervention relates to the choice of performance targets. Several approaches concerning performance target setting have been employed in different RBF programs (Oxman & Fretheim, 2009; Van Herck et al., 2010; Young et al., 2007). In RBF programs most often either absolute performance or relative performance targets have been used to gauge health provider performance (Young et al., 2007). Under the absolute performance targets health providers are rewarded when they achieve the pre-determined threshold of the target (e.g., 80% coverage of DPTs immunization among children under 5 years old). As for the relative-performance targets the high performers, for example those at the upper 90%, may be rewarded. Therefore, incorporating the performance targets in the health providers' contract can create an atmosphere of positive competition among health providers to drive performance and improve quality of care.

Despite the advantage of creating a sense of competition among health providers, these approaches are not free from considerable challenges during implementation. For example, the challenge with the absolute performance target is related to the existence of varying baselines in the target areas of different health providers – that is for the areas where the baseline is very low, it can be a challenging task for the health providers to reach the pre-determined target; on the other hand, for the areas with higher baseline, the health providers may not need to put much efforts to achieve the targets (Young et al., 2007). Likewise, a potential ceiling effect can create a serious implementation hurdle. The challenges associated with ceiling effects and the struggle of health providers to reach the targeted outcomes were also reported by Van Herck et al. (2010). In the case of ceiling effect, for example, it can be a real struggle to increase DPT3 coverage from 90% to 100% because vaccinating all eligible children for DPT3 in the target area will pose a solid implementation challenge. Similarly one should expect to encounter serious challenges with the relative performance targets as the

existing high performers may have an upper hand over the low performers. This imbalance can result in little efforts made by the high performers to achieve the targets due to the status quo (Young et al., 2007). It is also possible that some low performers do not participate in the performance-based program as they see their positions relatively weak and their chance of success very low (Young et al., 2007).

Performance indicators have been used in pay-for-performance contracts in some countries. Under the National Health Insurance program, for example, Belize implemented pay-for-performance contracts with public and private health facilities (Bowser et al., 2013). A larger portion of performance payments to health facilities was linked to the number of patients enrolled in the facilities each month. Similarly, performance indicators, instead of performance targets, were used in the contexts where no baseline figures were available for the purpose of contracting-out of health services. In Guatemala, for instance, due to lack of baseline values, performance indicators were used for contracting of health services (La Forgia, 2005). In the settings where no baseline is available, selection of performance indicators as a base for performance assessment of health providers can be considered a reasonable approach. However, the above approach may be accompanied with some serious drawbacks because assessing health provider performance on the basis of their progress against a list of indicators may not motivate them enough to improve performance, as there is a lack of adequate initiative to create a competitive environment among health providers to perform better and improve quality of care.

The structure and the way financial incentives are linked with the achievement of targeted services is another crucial aspect to take into account when a RBF program is designed. In Guatemala, linking a proportion of contract amount to performance was reported to be one of the key drivers in the ultimate success of the RBF interventions (La Forgia, 2005). It is common to use either financial withholds or performance bonuses for performance-based contracting in health sector (La Forgia, 2005; Custers et al., 2008). The concept of the two approaches is almost the same: paying extra amounts for achieving the planned targets. The main difference between financial withholds and performance bonuses is that in the financial withholds, a portion of a health provider's compensation is withheld by the purchaser until the health provider achieves the performance targets. In the performance bonus approach, a certain amount – extra to the contract budget – is retained by the purchaser and paid to the high performers or to those who achieve the predefined targets. For example, under the PPA

contracts, up to 1% of the total contract (extra to the contractual amount) was held by the Afghan MOPH, and paid to the high performers who had achieved 10 points in the mean score of all the 29 indices of the balanced scorecard (BSC) in the province (Hansen et al., 2008) . Despite the increasing use of the performance bonuses and financial withholds in many RBF programs, there exists limited empirical evidence on the relative effectiveness of either approaches (Custers et al., 2008).

A very relevant and critical question usually asked is *"what should be the size of incentives to motivate staff for better performance?"* In the literature, it is shown that between 5% and 10% of contract amount can be effective, particularly in the initial stages of the intervention (Young et al., 2007; La Forgia, 2005). Nevertheless, it is not unusual to see a much higher proportion of physicians' income from the P4P programs. For example, under the Quality and Outcome Framework in the UK, a performance payment which is an additional payment constitutes up to 25% of general practitioners' income (Custers et al., 2008; Campbell et al., 2009). Research shows a strong positive relationship between the amount of incentives and health provider performance (Basinga et al. 2011). However, there are also reports indicating no clear relationship between incentive size and the P4P results (Van Herck et al., 2010).

A dilution effect for incentive size has also been reported by several authors (Van Herck et al., 2010). Van Herck et al. (2010) stated that several studies in the USA showed a diluting effect for incentive size due to payer fragmentation which may have affected the P4P results. According to these studies, multiple payers who used different incentive schemes may have diluted the incentive size, and as a result there were fewer number of patients per provider, and lower incentive payments per health provider.

Communication with health providers and their awareness of P4P, and the way the program works, have also been reported as important factors that affect P4P results (Van Herck et al., 2010). According to the authors, several studies in the industrialized nations found a lack of or insufficient health provider awareness to be related to findings where there were no effects from P4P programs. Several studies found positive effects with those P4P programs that used extensive and direct communication with health providers involved in the program (Van Herck et al., 2010). Also, a study in the Netherlands reported positive results because of the involvement of health providers in designing the P4P programs (Kirschner et al., 2012).

With regard to the type and complexity of health services targeted, research shows that changing behaviour by paying incentives may be most difficult when providers treat patients with chronic and complex conditions due to increasing uncertainty over the patient health outcome (Town et al. 2004). On the other hand, it is easier to change behaviour by paying incentives when there is one-to-one relationship between an action and its desired outcome. A major problem of P4P program with the objective of treating patients with complex health conditions relates to serious unintended consequences. In the literature, several types of unintended consequences of P4P have been reported both in the developed nations and LMICs.

In the US, Roski et al. (2003), for example, studied the effect of a performance-based bonus with the objective of identifying patients with tobacco-use and providing tobacco cessation advice in large multiple speciality group practices. The authors found that documentation for tobacco cessation advice increased; however, provision of such advice for patients to quit smoking was not positively affected by the intervention. A similar problem was reported by Shen (2003) who examined the effect of a performance-based contracting program with the objective of increasing access to care for severely ill patients involved in substance abuse. The authors identified that fewer severely ill patients were treated by the program. The findings from both studies suggest cherry picking as a result of P4P, which means patients with mild to moderate health conditions were treated by the program, and patients with severe health conditions were ignored because of the likelihood of poor health outcome. Another unintended consequence of P4P relates to the problem of multitasking. This problem occurs when health providers distort their efforts and resources from the un-incentivised services to the incentivised services. In the US Rosenthal and Frank (2006), based on the evidence from their literature review on the effect of P4P, report that problems of gaming and multitasking were found as the unintended consequences of P4P in several programs.

In the LMIC countries, serious concerns were raised about the possible unintended consequences of RBF (Oxman & Fretheim, 2009; Kalk et al., 2010; Witter et al., 2012; Michael et al., 2013). In the literature review by Oxman & Fretheim (2009), the authors listed a list of undesired consequences of RBF – both from demand-side financing and supply-side financing – in the LMIC settings. With respect to the unintended consequences of demand side financing, the authors report on cases where mothers kept their children malnourished to be eligible for the Conditional Cash Transfer (CCT) programs in Mexico and Nicaragua. A

similar problem was reported concerning increased fertility rate from 2% to 4% as only pregnant women were eligible to benefit from the subsidies of the CCT programs. Therefore, the intervention instead of encouraging women to use family planning services, discouraged them to do so, as more women get pregnant to be eligible for the subsidies of the programs.

With regard to the possible adverse consequences of supply-side financing, Oxamn & Fretheim (2009) point out to the problems of distortions, gaming, corruption, cherry picking, dependency, demoralization, and dilution of intrinsic motivation among health providers. Distortions occur when health providers ignore their important tasks as a result of financial incentives. For example, they spend more time on documentation and reporting than on provision of patient care. Gaming occurs when health providers change their reporting rather than their actual provision of targeted services, for financial gains (e.g., over reporting of cases). Corruption in RBF might happen as financial incentives may be stolen or misused due to bureaucracy and mismanagement. Dependency may occur as health providers change behaviour because of financial incentives, and once the incentives are cut, their motives disappear. Demoralization and dilution of intrinsic motivation have a similar interpretation as the sense of dependency. Despite the growing concerns about the serious unintended consequences of RBF in LMICs, however, there is little research available to confirm the above concerns. In the developed countries, the gaming problem and adverse selection, which is a similar phenomenon as cherry picking, was reported in several studies (Petersen et al., 2006). However, Van Herck et al. (2010) reported that there was little evidence to suggest gaming problems in P4P programs in the industrialized countries.

With respect to ancillary support, which primarily refers to non-financial incentives, Oxam & Fretheim (2009) report that the use of financial incentives in RBF has commonly been part of a package that also includes non-financial incentives such as technical assistance, training, changes in management, and setting up of new information systems. The focus of both financial and non-financial incentives is to motivate health providers for better performance. The rest of the subsection discusses the literature about the determinants of health worker motivation, job satisfaction, and performance.

According to Franco et al. (2002), health worker motivation reflects the interactions between workers and their work environment in the context of respective culture and society. Because of the interactive nature of motivation, local organizational and broader sector policies have

the potential to affect motivation of health workers. Work motivation exists when achievement of organizational goals is associated with personal desired outcome such as a sense of personal achievement. In the motivation process, two interrelated psychological components work together: "will do" and "can do" components. The "will do" component refers to the degree to which a worker internalizes organizational goals. This component is dependent on the individual's work ethic and the intrinsic and extrinsic rewards emerging from the work. Some examples of intrinsic rewards are professional values and ethos, pride, recognition, and promotion. Some examples of extrinsic rewards are remuneration, benefits, salaries, and incentives. The "can do" component refers to the extent to which the worker effectively mobilizes his/her personal resources to achieve common goals of his/her own and those of the organization. This component is dependent on the worker's perception of his/her competencies and perception of availability of required resources and appropriate work environment (Franco et al., 2004). Determinants of worker motivation can affect one or both components, and this leads to the outcome of motivational process: worker behaviour and performance.

With the objective of identifying determinants of health worker motivation, in a study in four public hospitals in Jordon and Georgia, Franco et al. (2004) collected data in three phases: a contextual analysis, a qualitative assessment, and a quantitative in-depth analysis over a period of around ten months (between October 1999 and August 2000). The authors analysed the data in the context of both countries, and concluded that worker motivation is a complex concept affected by a whole range of determinants related to the worker, the organization where he/she works, and the cultural and societal environment where he/she lives and interacts with others. Therefore, effective interventions must operate on a set of key determinants, and need to address local cultural factors as well as broader societal factors that are affecting worker motivation at the local level. With respect to financial incentives, the authors state that financial incentives may be a major component, but on their own they risk having limited and even distorted impact on individual as well as on organizational performance.

Using the model developed and used by Franco et al. (2004), Mbindyo et al., (2009) conducted a study in eight district hospitals in Kenya with the aim of identifying influential factors on health worker motivation. The authors used qualitative method including in-depth interviews, small group interviews and focus group discussions with health workers from the

eight district hospitals. The authors emphasized on the important role of effective management and leadership at the hospital level in order to increase motivation of health workers. The health workers seemed to appreciate supportive leadership which may foster good working relationships between cadres, and address their expectations in terms of promotions, performance assessments, and good communications (Mbindyo et al., 2009). In another qualitative study conducted among health workers in Kenya and Benin, the role of human resource management tools and non-financial incentives was reported to be important for motivating health workers to adhere to their professional ethos, and to develop further qualifications (Mathauer & Imhoff, 2006).

The important role of human resource management and leadership has been emphasized in several other studies conducted in some African countries. In Tanzania, Manongi et al. (2006) conducted focus group discussions with health workers from primary health centres, located in 3 districts of a region in the country. The authors identified that health workers had a desire of more structured and supportive supervision from managers, improved transparency in career development opportunities, and more exchanges and communications between health facilities. In Uganda, Luboga et al. (2011) investigated determinants of physician motivation in 18 national hospitals, through several focus group discussions and administering job satisfaction questionnaires among physicians. The authors found that, after inadequate compensation, the largest contributors to dissatisfaction among physicians were poor quality of human resource management (e.g., staffing, workload, and professional development), and poor management of resources (e.g., availability of equipment, drugs, medical supplies, and quality of facility infrastructure). In Malawi, similar findings with regard to human resource management were reported by Chimwaza et al. (2014). Chimwaza et al. (2014) used a descriptive qualitative method, and examined influential factors among mid-level health workers. The most commonly cited critical factors were being treated unfairly or with disrespect, lack of recognition for their efforts, delays and inconsistencies in salary payment, and lack of transparency for promotion.

In a systemic review of 20 papers on determinants of health worker motivation, Willis-Shattuck et al. (2008) reported that within the context of cultural and societal settings, financial incentives, career development and management issues are core factors for health worker motivation. Financial incentives alone, however, are not sufficient to motivate health workers and retain them in the system. The authors reported that recognition is highly

influential, and that availability of resources and appropriate infrastructure are required to motivate health workers for better performance (Willis-Shattuck et al., 2008). With respect to the availability of resources and capacity for healthcare delivery in public health facilities, a study by Edward et al. (2011) shows significant improvement between 2004 and 2008 in Afghanistan. The authors analysed data collected annually from health facilities for the purpose of NHSPA,[10] and found that capacity for service provision – which included availability of essential drugs, equipment, clinical guidelines, and laboratory functionality, health provider adequacy, and provider knowledge – improved significantly over a period of five years. Using the same data source between 2004 and 2006, Edward et al. (2009) also reported significant association between availability of clinical guidelines and frequency of supervision and clinical quality of care provided from public health facilities. The above improvement in quality in terms of structure (e.g., availability of essential drugs, equipment, and clinical guidelines, etc.) has been as a result of contracting approaches in Afghanistan since the introduction of BPHS in early 2003.

With respect to the effect of contracting approach on clinical quality of care, Hansen et al. (2008) found positive results among the poor in Afghanistan. In a cross sectional study, using data from the NHSPA in 2004, Hansen et al. (2008) reported that overall, the quality of care at government-managed facilities and NGO-operated facilities did not differ statistically. However, the poor received significantly higher quality of care at NGO facilities than at government facilities. There was a close link between number of supervisory visits and quality of care. Training of lower level providers was significantly associated with clinical quality of care, defined as patient – provider interactions and communications.

Using the same NHSPA data source in Afghanistan, Edward et al. (2009) reported that between 2004 and 2006 quality of clinical care for sick children aged less than 5 years improved significantly. The authors found a 43.4% increase in the clinical assessment index, and a 28.7% increase in the counselling index over the three years. The clinical assessment index was strongly associated with certain characteristics of health providers (such as health providers being doctors, having a higher medical knowledge score, being trained in integrated management of childhood illnesses, and providing a longer consultation time).

[10] NHSPA stands for National Health Services Performance Assessment, which was a health facilities assessment program conducted annually in Afghanistan.

In Afghanistan, over a period of five years (2004-2008) significant improvement has been made in terms of improving clinical quality of care, capacity of service provision, patient satisfaction, and in terms of increasing equity in accessing healthcare (Edward et al., 2011). Using data collected through the NHSPA, the authors employed the generalized estimation equation models to assess the trends in the six domains of BSC between 2004 and 2008. The authors reported progressive improvement in the national median scores scaled from 0 – 100 in all the six domains of BSC. Except for the MOPH vision for pro-poor and pro-female health services, all changes in the domains were significant. The trends between 2004 and 2008 show significant improvements in the domains on patient and community satisfaction (65.3 – 84.5), health provider satisfaction (65.4 – 79.2), capacity of service provision (47.4 – 76.4), quality of services (40.5 – 67.4), financial management (84.4 – 95.7), and some improvement in the MOPH vision for pro-poor and pro-female health services (52.0 – 52.6).

Patient satisfaction and perceived quality of care in the context of PBF and contracting has been covered by several studies (Loevinsohn et al., 2009; Huntington et al., 2010; Soeters et al., 2001; Hansen et al., 2008). In the study by Huntington et al. (2010) in Egypt, the authors used a post intervention quasi-experimental design, and used data collected from exit interviews with patients. The authors reported significant improvement in patient satisfaction in terms of quality of family planning, antenatal care, and child health services in the facilities that received performance-based incentives. In Afghanistan Hansen et al. (2008) investigated the factors associated with patient perceived quality of primary care. The authors used data from the NHSPA, and identified that perceived quality was significantly higher when health providers were more through in taking patient histories and in conducting better physical examinations, and when health providers communicated better with patients. Patient perceived quality was significantly higher among the poorest quintile, and when the health provider was a doctor. However, perceived quality was significantly lower when patients spent an equivalent of US$ 2 or more for transportation to the health facility, and when patient waiting time in the facility was more than two hours (Hansen et al., 2008). While the above studies identified positive impact of contracting and PBF on patient satisfaction and perceived quality, a study in Uganda found adverse effects of PBF on patient satisfaction (Morgan 2010).

2.2 Patient satisfaction

The current literature offers very little empirical data on the impact of RBF and similar contracting approaches on patient satisfaction, and this is especially the case when it comes to the current RBF program in Afghanistan. It is widely believed that patient satisfaction is an indispensable element of care which can be affected to varying degrees by health provider performance, clinical quality of care, and by various attributes of health facilities. The remainder of the chapter reviews the literature concerning the evolution of customer satisfaction studies in healthcare in the industrialized nations, and the introduction of patient satisfaction surveys in the developing countries. The growing importance of patient satisfaction measures in relation to health outcome, and the rationale to include these measures into the performance-based contracting interventions will be discussed.

Ever since the first customer satisfaction studies in 1960s (Cardozo, 1965), there has been a proliferation in the publication of research on the subject with an estimated 15,000 academic and business articles published on consumer satisfaction over the past two decades (Peterson & Wilson, 1992). Since early 1980s customer satisfaction with healthcare has gained widespread recognition as a measure of quality in western countries (Newsome & Wright, 1999; Vuori, 1987; Williams, 1994). This is partly related to undertaking measures to engage patients in the healthcare process and partially due to the existence of strong relationships between patient satisfaction and patient compliance with treatment and medication use (Newsome & Wright, 1999).

The commonly cited conceptual model used in patient satisfaction surveys is 'disconfirmation theory' (Newsome & Wright, 1999; Oliver, 1980). This theory is based on the concept that satisfaction is a function of an initial standard (pre-purchased expectation) and some perceived discrepancy from the initial standard (Oliver, 1980). Disconfirmation theory proposes that a consumer compares his/her perception of a product or service against a pre-purchased comparison level or standard. Satisfaction is then mediated by the size and direction of the disconfirmation between the pre-purchased expectation and the experience with the performance or quality of the product or service. According to Oliver (1980), expectations are thought to create a frame of reference based on which one makes comparative judgement. In case of *negative disconfirmation* the outcomes are rated below the reference point, and in case of *positive disconfirmation* the outcomes are rated above the

reference point. The process of expectancy disconfirmation, however, is far from being a simplistic comparison between expectations and perceptions which highlights the fact that patient satisfaction in its widest sense is seen as a complex and multi-dimensional concept (Newsome & Wright, 1999).

While some empirical researchers have collected data on customer expectation and perception of a service, some others have raised concerns over the likelihood of serious biases occurring with the approach. An obvious example of such approach relates to the study by Mostafa (2005) who has compared the levels of expectation and perception of patients in a patient satisfaction study in eight Egyptian hospitals. The author found some evidence indicating that patients had higher levels of expectation in comparison with their levels of perception of quality care. With respect to such approach, Babakus and Boller (1992) argue that concurrent collection of data on the *expectation level* and *perception level* can result in potential bias because of response tendency that leads to higher levels of expectation being reported (a psychological constraint). Based on "disconfirmation theory" it is possible to measure patient satisfaction without collecting data concurrently on the expectation level and perception level because patients compare their perception of the quality of a service against a pre-purchased expectation. Cronin & Taylor. (1992) suggest that service quality can be predicted sufficiently by asking customers of their level of satisfaction or dissatisfaction rather than using the difference between the level of perception and level of expectation. This, however, does not mean that patient expectation is not important. On the contrary, asking patients about their level of satisfaction or dissatisfaction implies that their perception in comparison to their expectation of a service can be measured properly. The approach suggested by Cronin & Taylor (1992) has been used in the majority of empirical work on customer satisfaction, including patient satisfaction studies. In this PhD study, the above approach has been used for data collection on the levels of patient satisfaction with quality of care and performance of health providers and facilities.

With respect to patient expectation of health services in Afghanistan, a recent study conducted among Afghan parliamentarians deserves attention here. This study was conducted by SRTRO in 2010,[11] and its aim was to identify parliamentarians' perception and

[11] Silk Route Training and Research Organization (SRTRO) is a local NGO involved in monitoring and evaluation of health system in Afghanistan. SRTRO is currently carrying out the validation component of RBF program.

expectation of health services in Afghanistan (Silk Route Training & Research Organization, 2011). The SRTRO team found that only one quarter (25%) of the Afghan parliamentarians reported that people were happy with health services. This is despite the fact that nearly two thirds (65%) of parliamentarians reported that people in their communities have access to health services, and over two thirds (67%) of parliamentarians reported that quantity of health services are sufficient in the health facilities. But, the Afghan parliamentarians still think that people do not utilize health services because of concern over the quality of care. Poor quality of care was ranked second, after remoteness of health facilities, when parliamentarians were asked about the reasons for low utilization of health services. With respect to the poor quality of care, parliamentarians thought that the MOPH has not performed its regulatory role as much as people expected. Only 12% of parliamentarians said that MOPH was successful in regulating private pharmacies. Eighteen percent of parliamentarians stated that the MOPH was successful in regulating private hospitals. These suggest that majority (over 85%) of parliamentarians were either dissatisfied or unsure of the MOPH's capacity in terms of enforcing regulations and ensuring quality of care in the health sector, particularly in the private sector.[12] When it comes to the question as to whether the MOPH should continue working with NGOs, nearly 57% of parliamentarians supported the approach, and nearly 39% of them apposed the approach. All the above suggest that people have more expectations from the health system in Afghanistan.

Another crucial area with which Afghan people appear to be dissatisfied concerns corruption in public health facilities. In a recent study by Cockcroft et al. (2011), data were collected through a household survey from recent users of public health facilities in several districts of Kabul covered by the SM and a NGO. The authors found that 15% of service users in the SM facilities, and 26% of service users in the NGO facilities had to pay under-table fees to receive primary health care, which are free according to the Afghan constitution in public health facilities. The authors conducted focus group discussions with people in the catchment areas of the public health facilities. Focus group discussions confirmed that people knew payments were unofficial; but they were afraid to talk about corruption. The findings from Cockcroft et al. (2011) and from the study by SRTRO (2010) suggest that patient satisfaction, which is a reflection of patient expectation and quality of care, has to begin playing significant role in shaping the health system in Afghanistan.

[12] Private pharmacies and hospitals are not covered by contracting approach (by the SM or NGOs).

Using patient satisfaction measures as an evaluative device for healthcare quality has both its critics and advocates. According to the critics, using patient satisfaction as an instrument of assessing quality of care is often an unreliable approach. The critics argue that (i) patients may have a complex set of beliefs which cannot necessarily be expressed as satisfaction, and (ii) patients take technical competence of health providers for 'granted' due to the prevailing asymmetry of information (Williams, 1994).[13] Therefore, according to Williams (1994), patients assume a more passive role and may not often have expectations to evaluate providers' technical performance. The author further argues that patients often do not express dissatisfaction unless healthcare is extremely of poor quality. The same concern was raised by other critics as they found that patients often overrate quality of care (Peterson & Wilson, 1992). According to Peterson & Wilson (1992), patients may overrate quality of care due to social desirability issue and courtesy bias.[14]

On the other hand, advocates of patient satisfaction claim that patient satisfaction is a valid measure, and it reflects the side of quality that cannot be measured from the perspective of health professionals (Press & Fullam, 2011). Other advocates of patient satisfaction focuses on the ethical side of the argument, and lobby in favour of patient satisfaction surveys. For instance, Vuori (1987) argues that it is an ethical imperative to involve patients in quality assessment because health providers have monopoly of healthcare provision and their medical decisions and actions can affect patient health outcome. Vuori (1987) further argues that by including patient satisfaction measures, social accountability will be introduced to healthcare as the measures not only assure addressing patient needs, but also their perception of quality. Therefore, according to the author, patient satisfaction can be considered as a legitimate and desired outcome of care and a prerequisite for achieving the healthcare goals.

Some other advocates of patient satisfaction lobby for incorporating it in the payment scheme of RBF programs. Safavi (2006), for example, asks whether the current pay for performance initiatives work well for patient-centred healthcare delivery, as there is seemingly higher tendency towards technical performance in comparison with patient satisfaction.

[13] Asymmetry of information refers to the phenomena that a patient has less technical information and knowledge of his/her health problem than a health provider does(Gallouj, 1997).
[14] Courtesy bias occurs when a patient overrates quality of care because of respecting providers, and social desirability occurs when a patient overrates quality of care in order to avoid making providers feel unhappy.

While in developed countries advocates of patient satisfaction claim that it is beyond dispute that patient satisfaction is an essential measure of quality of care (Press & Fullam, 2011), in LMIC countries advocates of patient satisfaction contend that not enough attention has been paid to patient satisfaction (Rao et al., 2006; Andaleeb et al., 2007). For example, Andaleeb et al. (2007) argue that patients have very little voice in developing countries, and that patient voice has to begin playing a significant role in shaping the health system in LMICs. Other scholars such as Haddad et al. (1998) further support the notion of patient voice in defining the quality of care, and assert that only effective democratization can change the situation substantially and pave ground for active involvement of patients in devising healthcare system. Similarly, Rao et al. (2006) state that incorporating patient views into quality assessment offers one way of making health services more responsive to patients' needs.

The available number of studies, even though not many, on patient satisfaction highlights the growing interest in studying perceived quality of care in LMIC countries. In designing these studies, the theoretical models on patient satisfaction created in the developed nations have been adapted and used. The theoretical model commonly cited is the SERVQUAL framework, proposed by Parasuraman et al. (1988). The SERVQUAL framework proposes five dimensions for patient satisfaction: assurance, responsiveness, tangibility, empathy, and reliability. According to Parasuraman et al. (1991); Newsome & Wright (1999) "assurance" refers to the service provider's knowledge and courtesy and his/her ability to inspire customer's trust and confidence, "responsiveness" concerns the willingness of service providers to help customers and provide prompt service, "tangibility" refers to the appearance of physical facilities, equipment, personnel and materials, "empathy" denotes the service provider's caring and attention given to customers, and "reliability" refers to the ability of service providers to perform the promised service dependably and accurately. In patient satisfaction studies, the SERVQUAL framework has been used to assess validity of survey instrument. Validity in the context of patient satisfaction measurement refers to the ability of an instrument in capturing the five proposed dimensions of patient satisfaction. In my study, I use the SERVQUAL framework for the same purpose.

The SERVQUAL dimensions indicate that patient satisfaction can be influenced by a great number of factors related to attributes of health providers, and health facilities. Among these factors, behaviour of health providers has been identified as a significant predictor of patient satisfaction in several studies. For example, in a study in Bangladesh, Andaleeb et al. (2007)

identified that behaviour of nurses had the largest influence on patient satisfaction, followed by behaviour of doctors. Similarly several studies reported that health providers' behaviour were the strongest predictor of patient satisfaction in Bangladesh (Aldana et al., 2011; Andaleeb et al., 2007; Ashrafun & Uddin, 2011). Likewise, in an empirical work by Rao et al. (2006) it was revealed that doctor behaviour, and staff behaviour had the highest explanatory power on patient satisfaction in India. In addition, interpersonal interaction between a patient and his/her health provider has been identified as the strongest predictor of patient satisfaction in several other studies conducted in developing countries (Haddad et al., 1998; Baltussen et al., 2002; Abdulhadi et al., 2007; Hansen et al., 2008).

Similarly, according to several studies, factors such as availability of medicine, availability of medical equipment, and adequacy of resources had explanatory power, to varying degrees on patient satisfaction in LMIC countries (Haddad et al., 1998; Baltussen et al., 2002; Rao et al., 2006; Hansen et al., 2008). Also the influence of infrastructure, physical condition of building, tidiness of facility, and service capacity has been identified as significant, though not as strong as those of staff behaviour and interpersonal relationship (Haddad et al., 1998; Rao et al., 2006; Hansen et al., 2008; Atinga et al., 2011).

In most of the studies reviewed we have found that patients value more the humane aspect than the clinical aspect of healthcare delivery. The substantial role of communication skills and establishing good relationship and respecting patients in influencing patient satisfaction have been reported in several studies. The presence of various health resources has also been identified as significant factors in explaining patient satisfaction in LMICs' health context. The trends in improving patient satisfaction, and other aspects of healthcare suggest the existence of some relationships between patient satisfaction, and other aspects of healthcare. For the above relationships, two theoretical explanations can exist. First, patients may be satisfied because of better quality of care offered by the health facilities. This condition is supported by the evidence provided in the studies conducted by Baltussen et al. (2002); Hansen et al. (2008); Haddad et al. (1998); Atinga et al. (2011) . In the four studies in different LMIC countries, the authors identified that interpersonal interaction and communication skills of health providers are significant predictors of patient satisfaction.

Secondly, patients may be satisfied because improvement has taken place in other aspects of health care such as availability of various health resources as well as tidiness and appearance

of health facilities. The second explanation is further supported by the findings reported in the studies by Edward et al. (2011); Atinga et al. (2011). In the study by Atinga et al. (2011) in several Ghanaian hospitals, it was found that significant correlations existed between the hospital environment and appearance with patient satisfaction. Similarly a strong negative association was identified between waiting time and patient satisfaction.

With respect to performance-based financing, Meessen et al. (2011) argue that in RBF, health providers have the incentives to improve interpersonal relationships with patients to encourage higher utilization of services which ultimately increases their income. The authors assert that in a RBF program where the target recipients are health providers, it is likely that the health providers search for ways to attract more clients and maximize their income, and this might be achieved by improving quality of interpersonal communication. Furthermore, the authors state that health providers may be tempted to request their supervisors and managers for more health resources such as medicines, vaccines, equipment, clinical guidelines, and other supplies and equipment to improve quality of care with the aim of attracting more patients to maximize their income from the program.

As described in the earlier part of the chapter, the research on the effects of performance-based contracting on patient satisfaction provides some positive results. One of the studies which found positive impact of contacting of health services on patient satisfaction is the work by Loevinsohn et al. (2009). The authors reported significant increases in the levels of patient satisfaction with behaviour of health providers, with availability of medicines, and with tidiness of facilities over a period of two years in Pakistan.

The existing literature reviewed in this chapter shows that not enough work has been done to assess patient satisfaction in the context of RBF in LMIC countries, and in this respect the dearth of research is evident in Afghanistan. In this regard, the only study published so far is by Edward et al. (2011). However, the authors did not have an explicit objective of assessing the effect of contracting approaches on patient and community satisfaction. The authors were interested in assessing the performance of the healthcare system after the introduction of the BPHS which has been implemented through the performance-based contracting in Afghanistan. Even if the objective were to assess the impact of contracting, the cross sectional design with no control group used in the study cannot be considered a suitable approach for assessing treatment effects on quality of care and patient satisfaction.

Having a randomized controlled trial design, the current PhD study attempts to investigate the impact of RBF on patient satisfaction and its determinants at health facilities in Afghanistan. Moreover, this is the first study exploring the effectiveness of RBF by types of incentive administration mechanisms in a developing country. The findings from this study can contribute to both empirical circumstances and academic debates. In light of the study findings and contextual consideration, policy implications for the future of RBF will be provided for Afghanistan and other countries with similar settings. The evidence obtained from this study can contribute to the intense discussions surrounding the effects of RBF and similar approaches on quality of care and patient satisfaction in LMIC countries.

Chapter 3: Design and Implementation of RBF intervention

This chapter describes the structure of RBF incentives, the different incentive administration mechanisms, and the income of various types of health workers from RBF in Afghanistan. The chapter is organized into two sections. Section one focuses on describing the funding allocation, the performance targets and indicators, the amount of incentives allocated, and the validation of health facilities' reports. Section two describes the four types of incentive administration employed in health facilities. It also estimates health workers' income from RBF over the project period.

3.1 Incentive structure

The structure of incentives can be one of the most important steps in designing a RBF program. This is particularly the case in a post conflict situation such as Afghanistan where financial sustainability of the healthcare system depends on external funding aids. Continuity of donor funding may depend partially on the effective implementation of health interventions such as RBF and similar contracting approaches (Hecht et al., 2004).

3.1.1 Funding allocation

Of a total amount of US$ 12 million earmarked for RBF, the largest portion (around 85%) was allocated for RBF incentives targeted at health facilities. The remaining 15% was allocated for administrative and technical assistance which covered the PHOs' activities, the GCMU contract management, and the monitoring and evaluation activities of RBF program (Afghanistan Ministry of Public Health, , August 2009).

As pointed out in chapter one, the PHO team is one of the key stakeholders during RBF implementation in a province. They provide technical and managerial support to health facilities, and are the MOPH local authorities in the province. Their role in RBF is to streamline coordination at the province level amongst the various stakeholders, and monitor health facilities regularly. They receive performance incentives based on: (i) the number of facilities they monitor in a quarter; (ii) the number of minutes they record from the provincial health coordination committees (PHCC) in a quarter; and, (iii) the proportion of activities they have implemented from their quarterly work plan.

3.1.2 Performance targets and indicators

Under the RBF program in Afghanistan, health facilities are required to perform over the pre-determined targets in order to be rewarded the performance incentives. The pre-determined targets at the province level constituted a key part of the contractual obligation of an implementing partner (i.e. NGOs or SM) prior to the RBF intervention in a province. The existing contract was modified by the MOPH to add the RBF initiative. The initiative's objective was to encourage health facilities, managed by the implementing partner, to offer more of the target health services. The target health services were antenatal care, postnatal care, institutional deliveries, child immunization, and detection of tuberculosis cases.[15]

The MOPH used the reports from the HMIS to set the performance targets for each facility. The performance targets were determined for the RBF target services prior to the contract amendment, and prior to the randomization of health facilities. Once set, the performance targets became part of the contracts of RBF facilities in a province. Since target setting was carried out long before the randomization of health facilities, it is safe to say that targets were unrelated to treatment assignment, i.e. RBF status of health facilities.

The performance targets in the RBF program were used as threshold points, and health facilities were paid incentives if their outcomes exceeded the targets. Health facilities were required to report their progress against a list of output indicators. The indicators were number of antenatal care visits (ANC visit1, ANC visit2, ANC visit3, ANC visit4); number of postnatal care visits (PNC visit1, PNC visit2); number of deliveries conducted by skilled birth attendants (SBA), number of DPT3 vaccines administered to children, and number of tuberculosis cases detected in a health facility.

According to this arrangement, RBF health facilities were entitled to receiving incentives based on the extra number of cases reported over the targets every three months. For example, if a health facility conducted 90 cases of ANC visit1 in a particular quarter (or three months), and the target was 80 cases of ANC visit1, then, the facility was paid incentives for the 10 extra cases. No payment was made to those facilities whose report either was not validated by the third party or their performance was below the performance targets – no penalty was imposed on the latter either.

[15] These services are part of the BPHS, and have been offered from health facilities prior to the start of RBF

3.1.3 Incentive size

In the first year of RBF implementation, the incentives paid to health facilities constituted a small proportion of health providers' income. Our analysis of the available data shows that RBF incentives make up small percentage of health workers' income in the first year (Table 3.7). The percentages were 5.6% for salary-based, 6% for task-based, 3% for equal-amount, and 12% for mixed-method administration mechanisms.

The incentive amounts were determined based on unit cost by the MOPH. Based on the perceived workload and importance of each type of services, the MOPH came up with the incentive amounts per unit of target services. Initially incentive amount for each case of ANC visits and PNC visits was set in local currency at an equivalent of US$ 1.5. Each case of delivery conducted by a skilled birth attendance (SBA) was equivalent to US$ 12. The incentive for detecting a case of TB was equivalent to US$ 5, and for vaccinating each child with DPT3 was equivalent to US$3.

In the second year of RBF implementation, the MOPH decided to double the amounts for the ANC and PNC visits, and to increase the amount for each case of delivery by SBA to an equivalent of US$ 40. It is possible that the increase in the amounts for maternal health services was related to MOPH assessment that the initial amount was not sufficient to improve quality, and that maternal health services are high priority services for the ministry.

My analysis of the available data on the amounts paid to health facilities in the second and third years of RBF implementation indicates that incentives constituted sizeable portions of health workers' income. This was particularly true for the health providers involved in the delivery of maternal services. For example, in 2011 health workers' income from RBF incentives accounted for up to 18% in salary-based, 25% in task-based, 25% in equal-amount, and 18% in mixed-method administration mechanisms. In 2012, the proportion of health workers' income from RBF increased dramatically to 31% in salary-based, 38% in task-based, 32% in equal-amount, and 31% in mixed-method administration mechanisms (Table 3.7). The dramatic increases in the health workers' income from RBF might be related to the increased use of RBF target services.

3.1.4 Formula for computing incentives

Various RBF and pay for performance programs have adopted different formulas for processing the incentives. Some programs focus on the process, some on outputs, while some others on outcomes. It is not uncommon to see a hybrid form where the purchasers decide how much money to pay for the quantity and how much incentives to pay for the quality of healthcare (Basinga et al. 2011). In Afghanistan the amount of incentives that a facility receives, not only depends on the quantity but also on the quality of healthcare provided. In order to measure quantity the HMIS reports are used, and for the quality measures the scores obtained by the National Monitoring Checklist (NMC) are utilized.[16] The MOPH calculates the RBF incentives according to the formula below.

$$\text{Amount}_{it} = [\,\Sigma(P_j * U_{jit}) * \ Q_{it}] - 10\,\%$$

Where

Amount_{it} refers to the total amounts payable to facility *i* at time *t*

P_j stands for the amount paid per service unit *j*

U_{jit} refers to the number of patients receiving service *j* in facility *i* at time *t* over the target

Q_{it} displays the multiplier for overall quality of care for facility *i* at time *t*

10% refers to the 10% deduction of Amount_{it} payable to the NGO

The amounts of incentives are calculated based on the number of cases over the targets. The total amounts are, then multiplied by Q_{it}, which is a quality score produced by the PHO in the province.[17] The Q_{it} takes up to 100% and is computed based on quality of score measured by the use of NMC in each health facility. The NMC instrument is used to make sure the requirements of the BPHS are met. One of the key requirements, for example, is staffing level for which the tool measures the number and type of health workers, and provides score for the staffing domain according to the BPHS guideline. Similarly, based on the requirements of

[16] The NMC is a monitoring instrument developed by the MOPH to measure quality of care according to the BPHS standards.

[17] In the event a health facility is not monitored in a quarter due to security concerns, inaccessibility, blockage of road or any other reasons, the quality score from the previous quarter is used.

BPHS scores are calculated for all other domains (e.g., essential drugs, medical equipment, infection prevention, etc.). To produce the overall quality score, an average score of all the scores related to various domains in a health facility is computed. The overall quality score is, then used for computing the amounts of incentives the health facility is entitled to receive in that reporting quarter.

Ten percent of the incentives is given to the NGO (or SM) to cover the administration costs related to the RBF program. To illustrate how the total amounts of incentives for a facility was calculated, an example is provided from one of the facilities in Table 3.1. Table 3.1 shows the total incentive amount, based on the number of cases reported over the performance targets, is equivalent to US$ 768. The PHO reports a quality score of 75%. Therefore, the US$ 768 is multiplied by 0.75 and equals to US$ 576. Ten percent of US$ 576, which equals to US$ 58, is payable to the respective NGO for the RBF administrative cost. The remaining US$ 518 is payable to the facility to be distributed among health workers.

3.1.5 Validation of facility reports

In order to receive the incentives, the facility's HMIS reports related to the RBF target services had to be validated by third party (the Johns Hopkins University survey teams in Afghanistan). The purpose of validating the HMIS reports was to gauge the extent of discrepancy between the reports submitted by facilities and the reports prepared by third party. The third party established two mechanisms to validate the reports: (i) validation of reports at health facility level, and (ii) validation of reports at community level.

In order to carry out the validation of reports at facility level, the third party randomly selected 25% of the health facilities reports. The field monitors of the third party visited the sampled facilities. In line with the HMIS reporting period, the field monitors tallied the number of cases from the register books of the facilities and prepared validation reports. Then the reports were compiled in the Johns Hopkins University Office in Kabul, and compared with the reports that health facilities had submitted to the MOPH for the respective period.

For the validation of reports at community level, in each of the sampled facilities the third party field monitors randomly selected 10% of the patients who had utilized the RBF target services in the previous three months. The details of patients (e.g. names, addresses, or any other identifications) from a health facility's registers were written down in the designated form, and then given to the community monitors.[18] The community monitors interviewed women to find out whether the health services had been provided to patients. The community monitors compiled the reports and sent to the third party's office in Kabul. The third party analysed the reports, and sent the validation reports to the GCMU/MOPH. Discrepancies over 10% were considered a base for invalidating the health facility's reports. The facility was not paid any incentives, nor was it penalized for the mismatch of the reports.

3.1.6 Challenges associated with validation of reports

There were serious challenges when a facility's reports were validated at the community level. Major reasons for the mismatch of reports at the community level were wrong addresses and identification provided by patients (or caregivers), and the submission of validation reports to the third party's office in Kabul (Field Office of Johns Hopkins in Kabul Afghanistan, , November 2010). The first caveat created immense logistical problem for the community monitors who spent several hours in the field to identify the right patient who had visited the facility in question.19 The problem of recall bias by the patients (or caregivers) was another serious challenge. The problem of sending the community validation reports to Kabul added to the logistical burden, as the third party did not have field offices and it had to rely on the initiatives taken by the community monitors. The above arrangements may have increased the operation costs of validating facilities reports.

3.2 Incentive administration

The choice of payment mechanism can make a substantial difference in the ultimate success of a health intervention (La Forgia, 2005; Robyn et al., 2012). In RBF programs it is crucial to know how the incentives are paid to different categories of health workers in health facilities as this can significantly affect their performance and quality of care

[18] A community monitor was an educated woman selected from a community-based organizations and was trained to carry out the validation of reports at community level.
[19] This is particularly the case in rural areas of Afghanistan where there is no accurate home address, and where women due to cultural influences may not provide full identification while registering in health facilities.

3.2.1 Types of incentive administration

RBF incentives have been paid to health workers according to four administration mechanisms: salary-based, task-based, equal-amount, and mixed-method.[20] Structure of the four incentive administration mechanisms is described below.

Under the salary-based administration, the share of each health worker's incentive of the total incentive for the facility is proportionate to his/her salary. For example, if physician salary makes up 20% of the payroll for a health facility; then, a physician is paid 20% of the incentives the health facility receives in the respective quarterly period.

The task-based administration mechanism has been developed according to the level of involvement of a health worker in the delivery of RBF target services. For example, if a midwife manages all deliveries in a health facility, then the midwife receives the incentives related to institutional deliveries conducted in the facility. If two midwives manage deliveries, then the total for institutional delivery is split between the two.

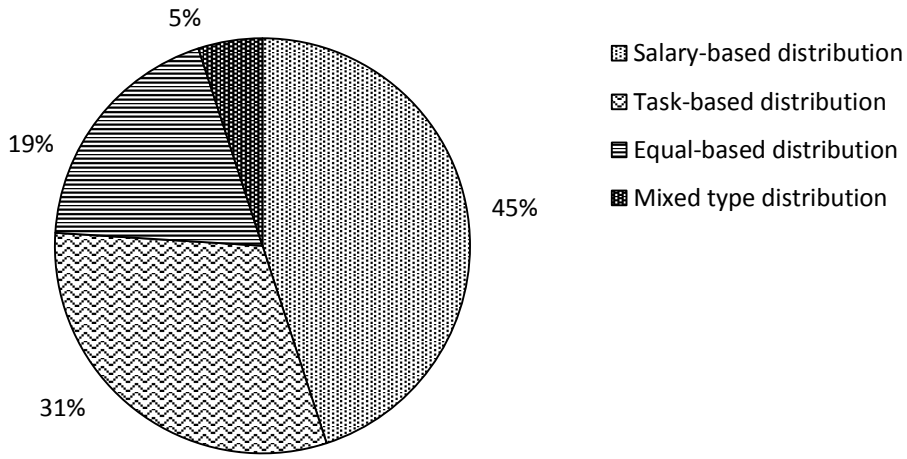
Under the equal-amount administration, the total amount of incentives a health facility receives is divided equally among the various categories of the facility staff, regardless of the salary status or position of the staff, or the extent of the staff's involvement in the delivery of RBF target services in the facility.

The forth type of incentive administration is the mixed-method mechanism which is a mixture of the salary-based and task-based administration types. In a health facility, the female health workers who are involved in the provision of antenatal care, postnatal care, and institutional deliveries are paid according to the task-based, and other staff are paid incentives according to the salary-based approach. Under this approach larger amounts are paid to female staff because of their exclusive engagement in the delivery of maternal health services, for which largest amounts of RBF incentives are allocated.

[20] Since the NGOs selected incentive administration types during project implementation, incentive administration mechanism is endogenous.

Percentage of RBF facilities employing the four types of incentive administration mechanisms varied in the 11 provinces (Figure 3.1). Figure 3.1 shows that less than half of RBF facilities (45%) employed the salary-based administration, nearly one third of them (31%) administered the task-based type, nearly one fifth (19 %) of the facilities adopted the equal-amount administration, and only 5% of the facilities used the mixed-method incentive administration mechanism.

Figure 3.1 RBF incentives by admin type



3.2.2 Estimation of salary package for health workers

Income of health workers is made up of their salaries, night duty wage,[21] hardship allowance, and performance incentives - herein after we use the term salary package for these sources of income.[22] I estimate the salary package by using the MOPH National Salary Policy (Afghanistan Ministry of Public Health, , December 2011). The estimated salary package will be used as the denominator when I calculate the proportion of health workers' income from the program. The following subsection describes the way the salary package is estimated.

[21] During weekdays, night duty starts from 4:00pm on a day and lasts until 8:00am on the following day. During weekends, it covers 24-hour services offered by the on-duty health workers.

[22] In Afghanistan health providers also earn from other sources such as their private clinics and potentially from non-official charges or under table payments in health facilities (Cockcroft et al., 2011) . It is very difficult, if not impossible, to collect accurate data on the income of health workers in Afghanistan.

In the MOPH National Salary Policy (NSP), the base salaries and night duty allowances are determined according to the type and sex of health workers. Similarly, the hardship allowances are specified according to the geographic location of health facilities. Table 3.2 shows that there are five categories of health facilities according to the geographic locations: (i) urban, (ii) semi-urban, (iii) rural, (iv) deep rural, and (v) isolated areas. Table 3.2 also shows the number of RBF facilities located in each type of geographic category. It is clear from these data that the majority of RBF facilities are located in rural, deep rural, and isolated areas, and for these facilities the amounts paid as hardship allowances are substantially higher than those paid to facilities located in semi-urban area. All categories, except for the health facilities located in urban area qualify for receiving the hardship allowances.[23]

The MOPH NSP specifies certain criteria to determine the hardship allowance category of a health facility. For example, distance from the provincial capital, girls' enrolment rate in the primary school in the area,[24] and several other criteria are used to determine the scores for computing the hardship allowances (Afghanistan Ministry of Public Health, , December 2011). The higher the scores, the greater the hardship allowances. For instance, a facility with a total score of 91-100 qualifies for the category of isolated area, while another facility with a total score of 25-40 qualifies for the category of semi-urban area.

In addition to the geographic hardship allowances, health providers are paid for doing night shifts. According to the MOPH national salary policy, 500 Afghani (equivalent to $US 10) per night is paid for a doctor, 400 Afghani (equivalent to $US 8) per night per person is payable for other medical staff, and 300 Afghani (equivalent to $US 6) per night per person is paid for support staff as night duty allowances. I estimated a total of 10 night duties per person for the categories of doctor, nurse, midwife, and assistant midwife (Table 3.3).

Table 3.3 presents an example of how salary package works. The table demonstrates the base salaries, night duty allowances, hardship allowance and RBF incentives which sum up to the salary package for different categories of health workers working in a CHC located in a semi-urban area. The monthly base salary for a midwife is US$ 178, the hardship and night duty

[23] Hardship allowances are used in developing countries to attract and retain health workers in remote areas (Henderson et al. 2008).
[24] Girls' enrolment rate has been used as a proxy for classifying a geographic location as rural, semi-urban, and urban area, because in rural areas girls' enrolment rate in schools is much lower than other geographic areas.

allowances are US$ 89 and US$ 80 respectively, and RBF incentive is US$100 which all sum up to US$ 447. The MOPH NSP also recommends an increase of 5% of the base salaries to account for annual inflation rate. I also employ this recommendation in my estimations.

3.2.3 Health workers' income from RBF

For the *salary-based administration*, first the proportion of each category of health worker salary out of the total payroll was calculated to provide a weight for the category. The weight was multiplied by the total incentive to obtain the amount the category has received.

The above process is illustrated in Table 3.4. In Table 3.4 a hypothetical scenario has been portrayed where the total payroll amount is U$2400 and the share of each category of health workers' salaries is estimated by diving the total salary for the category by the total payroll amount. The total incentive for the facility is multiplied by the share of each type of health worker's salary to determine the amount paid to the health worker. For example, if the doctor's salary is U$421, then his/her share of the payroll amount (U$2400) is 0.18. If the total incentive amount paid to the health facility is U$600, then the doctor's share of this amount is U$105 (=US$ 600 x 0.18).

For the *task-based administration* mechanism weights were calculated for ANC visits, PNC visits, institutional deliveries, DPT3 coverage, and TB case detection by using the number of cases of each type of cases exceeding the performance targets.[25] Since the MOPH increased incentive amounts for ANC and PNC visits, and for institutional deliveries in the second year, two types of weight – before the increase and after the increase in the incentive amounts were computed, and then average weights for each type of services taken.

In order to estimate the weights in the first year, for all RBF facilities the number of cases over the performance targets was multiplied by the incentive amount assigned for the category of service. For example, if the HMIS reports showed that ANC and PNC visits were 28,300 cases over the performance target, then total amount for ANC and PNC visits was US$ 42,450 (=28,300 x US$ 1.5), if institutional deliveries were 3,300 cases over the performance target, then total amount for deliveries equalled to US$ 39,600 (=3,300 x US$ 12), if DPT3 vaccination was 5,800 cases over the performance target, then total amount

[25] To ensure accuracy of the reports, the reports which had been verified by the third party were used.

for DPT3 equalled to US$ 17,400 (=5,800 x US$ 3), and if number of TB cases detected over the performance target was 150 cases, then total amount for TB cases detected equalled to US$ 750 (=150 x US$ 5). Total amount for each category of service was divided by the grant total amount in order to get the weight for the category. For example, weight for DPT3 vaccination was 0.17 (= 17,400 ÷ 100,200), as grant total amount was US$ 100,200 (= 42450 + 39600 + 17400 + 750). The same approach was used for other categories of services.

The same approach was used for computing weighs in the second year, and then average weight for the first and second year was taken for each category of services. The weights for ANC visits, PNC visits, and institutional deliveries were added to be used as weight for maternal services.[26] The weights were: for maternal services 0.82 (= 0.31 for ANC + 0.12 for PNC + 0.39 for deliveries), for DPT3 vaccination 0.17, and for the TB case detection 0.03.

To calculate the incentive amounts, I multiplied the total incentives paid to a health facility by the weights for maternal services, DPT3 vaccination, and TB case detection. The incentive amounts were split between those staff members involved in the provision of the service. An illustration of how the task-based mechanism works is displayed in Table 3.5. Table 3.5 shows that a health facility receives US$ 1632 in a quarter. Of this amount, US$ 1312 (= US$ 1632 x 0.82) is allocated for maternal services, US$ 272 (=US$ 1632 x 0.17) for DPT3 vaccination, and US$ 48 (=US$ 1632 x 0.03) for TB case detection. The US$ 1312 is then divided among female doctor, midwife, community midwife, and female nurse as each of them receives US$ 328. Likewise, the amount for DPT3 coverage is split between vaccinators.

For the *equal-amount administration*, I divided the incentive amounts a health facility received by the number of health workers. In this way, I obtained the amounts each health worker received from the RBF incentives. Then, I divided the amount each category of health workers received by the salary package, and obtained the proportion of incentives the health workers have received. An illustration of the equal-amount administration is provided in Table 3.6. Table 3.6 shows that each type of facility staff receive the same amount of incentive.

[26] The three services are provided by the same type of health providers, therefore, the weight applies to their incentives.

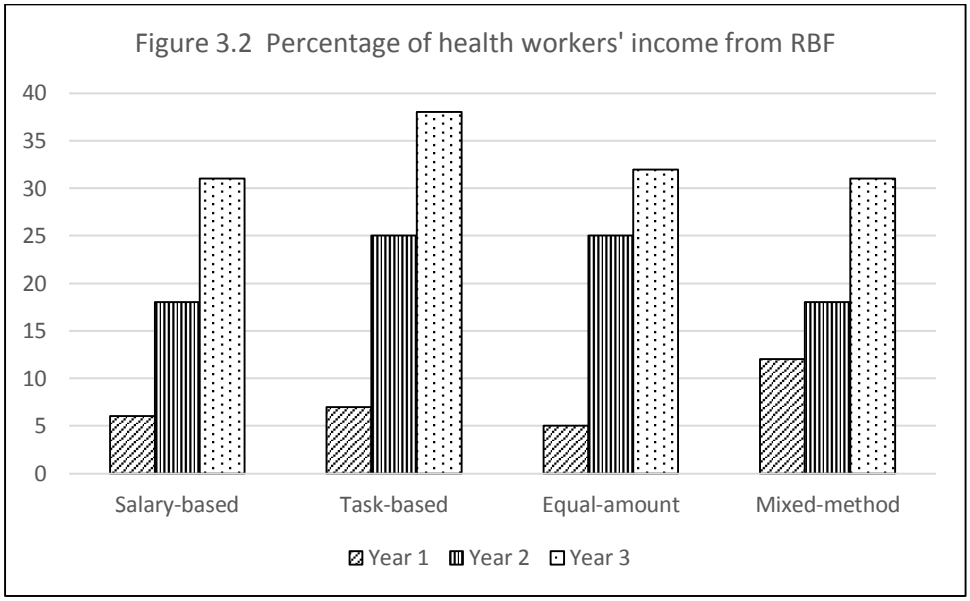
For the *mixed-method administration*, both the salary-based and task-based approaches were employed in a health facility. The task-based approach was used for distributing financial incentives among those female health workers who were directly involved in conducting the institutional deliveries, and the salary-based method was employed for distributing monetary incentives among other types of health workers.

3.2.4 Proportion of health workers' income from RBF

In each of the project implementation year, proportions of health workers' income from RBF were obtained by dividing the amounts of incentives health workers received by the amounts of salary package (Table 3.7).[27] The findings in Table 3.7 are interpreted in the context of Afghanistan. The proportion of health workers' income from RBF are quite small in the first year: 6% in the salary-based, 7% in the task-based, 5% in the equal-amount, and 12% in the mixed-method administration mechanisms. In the second and third years, particularly in the third year the percentages increased (Figure 3.2).

Figure 3.2 shows that in the second year for those who received salary-based incentives income from RBF was 18% of total income. For the health providers who received task-based incentives and equal-amount incentives, income from RBF was 25% of total income for each, and for those facility staff who received mixed-method incentives income from RBF was 18% of total income. In the third year, proportion of health workers' income from RBF increased substantially. Of total income, health workers' income from RBF was 31% under the salary-based, 38% under the task-based, 32% under the equal-amount, and 31% under the mixed-method administration mechanisms. The increase in the health workers' income from RBF over time suggests that more target services were provided in the second and third years, and that the amount of incentives increased substantially for ANC visits, PNC visits, and institutional deliveries.[28]

[27] For the third year, providers' income from RBF was divided by salary package of seven months because their income was from September 2012 to March 2013.
[28] First year refers to September 2010 – August 2011, second year refers to September 2011 – August 2012, and third year refers to September 2012 – March 2013.

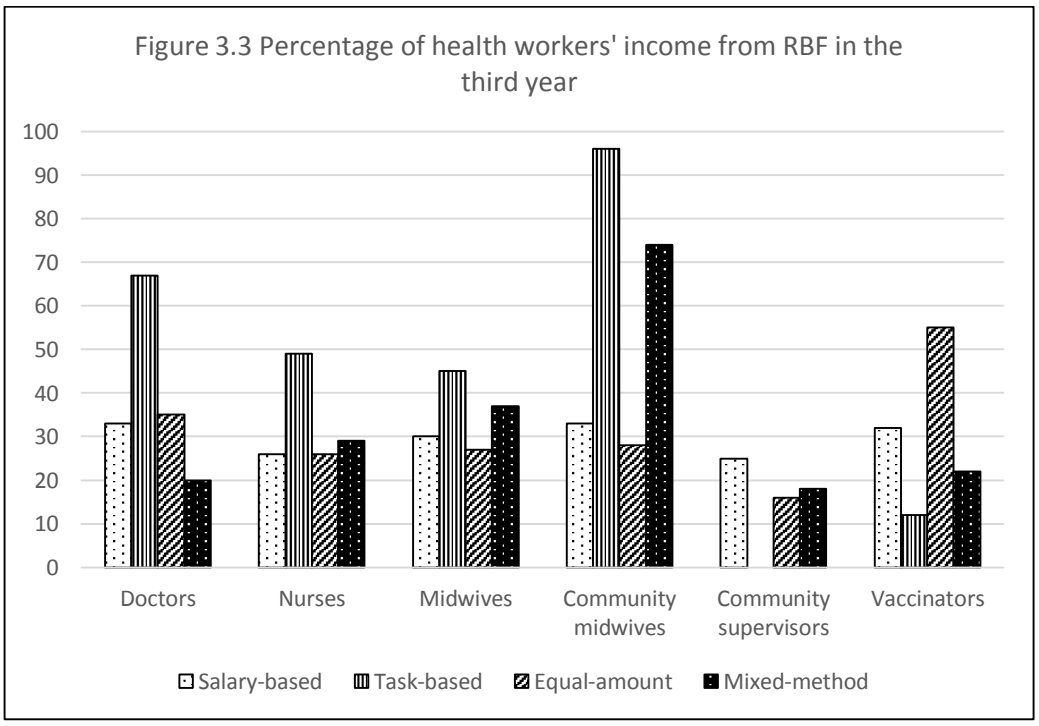


Considerable differences existed in the proportions of health workers' income between the four types of administration in the second and third years. The difference was more obvious between the task-based versus the other three administration mechanisms. Under the task-based administration, the proportion of health workers' income who were involved in the delivery of maternal healthcare were almost two to three times the incentives of health workers in the other three administration types. For example, in the second year, of their total income, community midwives' income from RBF was 61% in the task-based, 20% in the salary-based, 23% in equal-amount, and 33% in the mixed-method administration mechanisms. The above differences might be due to the larger amounts paid for the delivery of maternal healthcare to the health workers under the task-based versus other types of administration.

On the contrary, under the equal-amount administration the proportions of health workers' income who were involved in the immunization program and administrative work were more than twice of the health workers' income in the other three administrations. For example, in the second year the vaccinators' income out of total income was 42% under the equal-amount, 20% under the salary-based, 7% under the task-based, and 13% under the mixed-method administrations. This may be related to the smaller sizes of base salaries for vaccinators.

Similarly, considerable differences were seen in the proportions of health workers' income between the various categories of health staff within each type of administration in the second and third years. Here again the task-based and equal-amount administrations experienced large differences in the health workers' income. For example, under the task-based approach in the second year the proportions of community midwives' income from RBF was 61% and of vaccinators' income was 7%. In the third year, they were 96% for community midwives and 12% for vaccinators.

It is of interest to compare the proportions of each category of health workers' income from RBF under the four incentive administration types. Figure 3.3 shows that health workers' income from RBF increased largely in the third year. The doctors' income was 33% under the salary-based, 67% under the task-based, 35% under the equal-amount, and 20% under the mixed-method administration mechanisms. There was a huge variation in the income of community midwives from RBF. For example, share of community midwives' income were 33% in salary-based, 96% in task-based, 28% in equal-amount, and 74% in mixed-method administration mechanisms. From the above it becomes clear that those health workers directly involved in the delivery of maternal health services, particularly the midwives and community midwives earned larger amounts from the RBF intervention.



The difference was higher under the task-based versus other administration types. A very different pattern existed for the community supervisors, and this may signal the very little attention paid towards the community-based activities. The community-based activities are mainly provided by community health workers (CHWs) who are supervised by community health supervisors.

3.2.5 Involvement of community health workers in RBF

Another serious concern relates to the limited attention paid to CHWs who are not paid any incentives, but who play crucial role in the BPHS program. The CHWs constitute the community-based structure of the BPHS (Newbrander et al., 2014), and this structure is the key linkage between the community and health facilities. CHWs are volunteers who provide essential maternal and child health services through health posts located in villages at the catchment areas of health facilities. CHWs refer patients to health facilities and do follow up visits for certain types of services, including the postpartum patients.

Except for the health facilities under the mixed-method administration, the role of CHWs has not been taken seriously into account as no incentives in any forms were considered for them in the RBF program. This issue is crucial as the RBF intervention could have been a good opportunity to allocate some funding to CHWs' related activities, and ask the implementing NGOs to come up with initiative plans to pay incentives in some forms to these volunteers and increase the utilization of maternal health services.

3.3 Conclusions

A comparison of my findings about health workers' income from RBF with other studies indicates that the above percentages are consistent with the varying percentages of performance incentives (ranging from 5% to 25% of providers' income) employed in the health sector of many countries. It becomes obvious that no consensus exists among scholars on the size of monetary incentives. In most RBF and similar programs, between 5 to 10% of the provider's income are reported to be sufficient to motivate health providers for better performance (La Forgia, 2005; Young et al., 2007; Peabody et al., 2011). With respect to RBF in Afghanistan, the proportion of health provider's income from RBF seems to be in line with the commonly practiced ranges; but, with varying degrees among different categories of health workers.

In my study, it was found that certain types of health workers' income from RBF increased substantially over time; however, considering the poor purchasing power of Afghani currency even those seemingly attractive amounts may not be sufficient to motivate health providers for better performance unless structural and organizational elements (e.g., non-financial incentives, essential infrastructure, etc.) are sufficiently provided to health facilities.

Another important consideration relates to the scope for changing provider behaviour by paying incentives. Research shows that changing behaviours by paying incentives may be most difficult when providers treat patients with chronic and complex conditions due to increasing uncertainty over the patient health outcome (Town et al. 2004). Unlike the above, in the RBF program in Afghanistan the relationship between each target service and the incentive is fairly straightforward: a pre-determined amount is payable for each extra case of RBF target services to a health facility after the verification of a health facility's reports.

Table 3.1 Incentives computed for one of the CHCs in the third quarter of 1391 (Solar calendar)

Target service	Performance targets	Number of cases reported	Cases over the performance target	Amount per case in US$	Amounts in US$
ANC visit1	225	172	0	1.5	0
ANC visit2	39	102	63	1.5	94
ANC visit3	9	38	29	1.5	43
ANC visit4	6	10	4	1.5	6
PNC visit1	99	83	0	1.5	0
PNC visit2	9	67	58	1.5	87
Deliveries by SBAs	48	83	35	12	420
DPT3 vaccines	144	183	39	3	117
TB cases detected	0	0	0	5	0
Total amount					768
Quality score for the facility					75%
Total amount after multiplied by the quality score					576
10% for the NGO					58
Amount payable to the facility					518

Note: This facility was paid slightly over US$ 468 instead of the amount calculated here (US$ 518)

Table 3.2 Hardship allowances in various geographic locations

Geographic location	Number of RBF facilities	Hardship allowance	
		Male	Female
Urban	13	0%	0%
Semi-urban	35	25%	50%
Rural	80	50%	100%
Deep rural	80	100%	200%
Isolated	22	125%	250%
Total	230		

Table 3.3 An example of monthly salary package for health workers in a semi-urban CHC

Position	Base salary in US$	Hardship allowance In US$	Night duty allowance In US$	RBF incentive In US$	Salary package In US$
Doctor (male)	214	107	100	80	501
Nurse	138	69	80	60	347
Midwife	178	89	80	100	447
Community midwife	126	63	80	120	389
Pharmacist	190	95			285
Vaccinator (male)	104	52		40	196
Community supervisor	126	63		10	199
Administrator	126	63			189
Guard/cleaner	100		60		160

Night duty allowances are estimated based on 10 nights per month

Table 3.4 An hypothetical example of salary-based administration

Payroll amount	Position	Salary	Weight of salary based on payroll	Incentive for the facility	Incentive per person
2400	Doctor	421	0.18	600	105
	Midwife	347	0.14		87
	Com-midwife	269	0.11		67
	Nurse	287	0.12		72
	Vaccinator	156	0.07		39
	Guard/cleaner	160	0.07		40

The amounts are shown in equivalent of US$; though, the staff were paid in local currency

Table 3.5 An hypothetical illustration of the task-based administration

Health services	Weight	Incentive per service type	The health workers involved in the task	Amount paid to health workers
Maternal services (ANC and PNC visits, deliveries)	0.82	1312	Female doctor	328
			Midwife	328
			Community midwife	328
			Female nurse	328
DPT3	0.17	272	Vaccinator fix-centre	136
			Vaccinator outreach	136
TB case finding	0.03	48	Male nurse	48
Total amount	1	1632		1632

Table 3.6 An illustration of the equal-amount administration

Health facility staff	Amount paid to staff in US$
Female doctor	320
Male doctor	320
Midwife	320
Community midwife	320
Female nurse	320
Vaccinator fix-centre	320
Vaccinator outreach	320
Administrator	320
Guard	320
Cleaner	320
Total 10 staff	3200 US$

Table 3.7 Proportion of health providers' income from RBF under the four administration types

Position	Salary-based				Task-based				Equal-amount				Mixed-method			
	N	Year 1	Year 2	Year 3	N	Year 1	Year 2	Year 3	N	Year 1	Year 2	Year 3	N	Year 1	Year 2	Year 3
Doctor	81	6	19	33	5	6	36	67	29	5	28	35	13	10	13	20
Nurse	114	6	18	26	9	10	47	49	44	4	22	26	6	14	18	29
Midwife	82	6	20	30	72	7	25	45	31	4	22	27	17	18	26	37
Community midwife	54	7	20	33	22	17	61	96	24	3	23	28	1	22	33	74
Community supervisor	47	5	14	25					14	3	13	16	3	4	8	18
Vaccinator	139	6	20	32	103	5	7	12	67	7	42	55	12	9	13	22
Others	37	5	18	37	77	0.7	0.2	0.3	15	6	26	35	2	10	13	17
All positions	554	6	18	31	288	6.5	25	38	224	4.5	25	32	54	12	18	31

In the task-based administration, the categories of doctors, nurses, midwives and community midwives refer to the female health workers

Chapter 4: Study design and methodology

This chapter first discusses the conceptual framework which has largely benefited from the literature review in chapter 2. Then it describes the hypothesis, study design and sampling approaches, and the data collection process. In the last part of the chapter, the estimation model, choice of regression models, choice of outcome and control variables, and the measures to address clustering effects are discussed.

4.1 Conceptual framework

In chapter 2, we have found that PBF programs have largely been part of bigger interventions which included not only financial incentives but also various non-financial initiatives aiming to improving performance of the health sector in terms of service utilization, quality, equity, and efficiency which are directly mediated by the health worker motivation to apply themselves to their tasks. Worker motivation can be defined as an individual's degree of willingness to exert and maintain an effort towards achieving organizational goals (Franco et al., 2002). Worker motivation is a complex and internal process which crosses many disciplines, including psychology, economics, social, organizational development, and human resource management.

Relevant to the discussion is the conceptual framework crafted by Franco et al. (2002). According to Franco et al. (2002), there are three types of determinants that influence health worker motivation: the internal individual level determinants, the organizational level determinants which refer to the work context, and the broader societal culture determinants. The interactions between individuals and their work environment, and the fit between these determinants with those of the broader societal culture, determine health worker motivation.

In the context of RBF, health worker motivation can be affected by the intervention where organizational culture, reporting structure, human resource management, units and channels of accountability, type of interaction with clients, and level of health workers' efforts exerted might be affected. Health worker motivation might be influenced by financial incentives, but they are not sufficient to address all problems related to low motivation. At the organization level, work environment, including availability of resources and a supportive management system, coupled with required level of competency and communication skills of workers, in addition to adequate monetary pays, can affect worker motivation for better performance. In

addition to these factors, the individual level determinants that stimulate a worker to come to work regularly, work diligently, be flexible and be willing to carry out the assigned tasks are crucial. According to Franco et al. (2002), in addition to organizational level determinants, clients and community members can also influence health worker motivation through their expectations for how services should be delivered, through their interactions with health workers, and through formal and informal feedback they provide on health worker performance.

Another aspect that we have identified in chapter 2 which was, to some extent, affected by RBF, concerns quality of care. Quality of care is a multidimensional concept. There are many definitions of quality of care depending on specific application of a health program (Penneys, 1997). In order to be able to assess RBF effect on some aspects of quality of care, I use the well-known quality framework developed by Donabedian (1988). According to the Donabedian framework, information about quality can be obtained from three categories of attributes: "structure", "process", and "outcome". Structure refers to the attributes of the health care settings where care is provided. Structure includes the attributes on material resources (e.g., facility building, equipment, medicine, vaccines, clinical guidelines, and financing), human resources (e.g., the number and types of health workers), and the organizational structure (e.g., availability of job descriptions, reporting system, medical records, methods of payment, and management and supervision system).

Process denotes to what is performed during the delivery of care to patients, and how it is done. For example, what activities a health provider performs in assessing patients and making diagnoses, recommending treatments, and providing medical advice to patients. Outcome measures attempt to describe the effects of care on the health status of patients and populations. Outcomes in healthcare can be expressed as: death, diseases, and satisfaction or dissatisfaction. Diseases refer to symptoms which are expressed by patients (e.g. pains, breathing difficulties, etc.), physical signs which are mainly discovered by health providers, and disabilities are physical or mental health condition(s) which impairs a patient's daily routines at home, at work or in recreation. Satisfaction or dissatisfaction are emotional reactions to diseases, and to the attributes of structure, process of healthcare provision, and outcome of healthcare delivered. Both structure and procedure can affect the outcome. It also possible that one outcome be considered as a process for another outcome. For example, taking patient satisfaction as an outcome measure, a patient might have a happy feeling

because of the availability of medicine prescribed by the health provider (structure), or because of the way the health provider has communicated with her/his during their visit (process), or because the treatment which was recommended by the health provider was effective and the patient has recovered from an illness (outcome). In another instance, another patient might have an unhappy feeling because the medicine prescribed is not available in the facility (structure), or because the patient's expectations of privacy during her/his visit with the health provider have not been taken respected (process), or because and the pain killer administered by the health provider has not relieved the patient's pain to the degree she/he has expected (outcome). These feelings and emotional reactions can be expressed by patients as their levels of satisfaction or dissatisfaction with quality of care provided by a health provider or facility.

With respect to determinants of patient satisfaction, in chapter 2 I have reported findings from several studies conducted in developing countries. Some of these studies have identified health worker behaviour as strong determinant of patient satisfaction (Rao et al, 2006; Andaleeb et al., 2007; Aldana et al., 2010; Ashrafun & Uddin., 2011). Some others have reported communication skills of providers to be strong predictor of patient satisfaction (Haddad et al., 1998; Baltussen et al., 2002; Abdulhadi et al., 2007; Hansen et al., 2008).

Availability of resources, tidiness and appearance of health facility and its attributes, and waiting time have also been reported as strong predictors of patient satisfaction, though not as strong as communication skills and behaviour of health providers. In the studies by (Haddad et al., 1998; Baltussen et al., 2002; Rao et al., 2006; Hansen et al., 2008) it was reported that availability of medicine, medical equipment, and other resources can affect patient satisfaction at varying degrees. Similarly, facility building, tidiness of infrastructure, physical appearance, and capacity for health provision of health facilities have been reported as significant predictors of patient satisfaction (Haddad et al., 1998; Rao et al., 2006; Hansen et al., 2008; Atinga et al., 2011). A negative and strong correlation was also reported between waiting time and patient satisfaction (Atinga et al., 2011).

In the context of RBF, it was reported that health worker behaviour have had significant effect on patient satisfaction (Loevinsohn et al., 2009). In the same study, the role of availability of medicine and tidiness of health facilities has also been reported as significant. With respect to behaviour of health providers, Meessen et al. (2011) argue that in the context

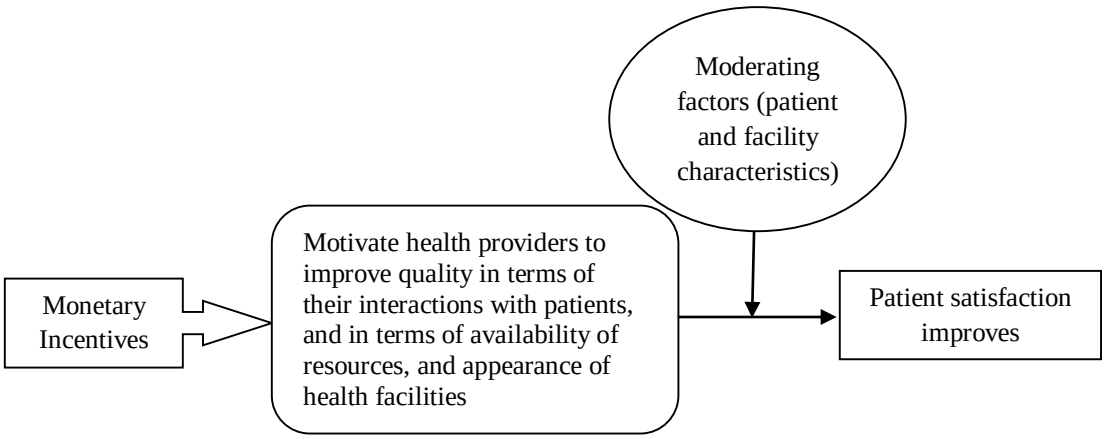
of RBF, health providers may change behaviour to improve perform because of the incentives. They may provide quality care to satisfy patients and attract more service users to maximize their income. Health providers may put efforts to improve communication skills and build rapport and trust with patients. They may spend more time examining and treating patients, and the patients may be satisfied with the way health providers perform and/or the way health providers communicate with patients (Meessen et al., 2011).

Other scholars, however, assert that in addition to financial payments, availability and organization of various resources, coupled with clinical competency and skills of health providers are crucial to meet patient expectations in terms of quality care, and provide health services to their satisfaction. In this regard, the findings from (Edward et al., 2009; Edward et al., 2011; Chimwaza et al., 2014) are relevant to mention. In these studies significant relationships between availability and organization of resources and quality of care are reported in two developing countries.

Based on the design of the RBF in Afghanistan, I are interested in investigating as whether rewarding health facilities with financial incentives improves quality of care that results in higher levels of patient satisfaction. Likewise, I examine the impact of RBF on structural and organizational elements of quality, and on health provider performance because these factors can affect patient satisfaction significantly (Andaleeb, 2008; Andaleeb et al., 2007; Atinga et al., 2011; Haddad et al., 1998; Rao et al., 2006).

In light of the discussion presented so far, and considering the design of RBF in Afghanistan I use the principal – agent theory to examine the intervention's effect on patient satisfaction. Principal – agent theory suggests that by linking incentives to the achievement of performance outcomes specified in a contract can motivate the agent to achieve the outcomes as desired by the principal. In the current RBF context, the principal is the MOPH, and the agent refers to health facilities, managed by NGOs. As can be seen, there is no direct contracting relationship between the MOPH and health facilities as this relationship is mediated by NGOs. Each NGO in its own right acts as the principal, and maintains its relationship with health workers through an employment contracting arrangement.

According to the design of RBF in Afghanistan, it is expected that paying financial incentives to health workers who are employees of NGOs can improve quality of care and increase quantity of targeted services through motivating health workers. The improvement in quality of care and health provider performance is expected to improve patient satisfaction with care in public health facilities. The relationships between financial incentives with potential of improving health provider performance and quality of care which can affect patient satisfaction are depicted in Figure 4.1 below.



Hypothesis: Paying monetary incentives, in addition to regular salaries, to health providers:

1. Positively affects patient satisfaction.
2. Positively affects health provider performance.
3. Positively affects availability of medicine, vaccines, and equipment in health facilities.
4. Positively affects physical condition and cleanliness of health facilities.

As described in chapter 1, in Afghanistan the RBF incentives were distributed among health providers in health facilities according to four types of incentive administration mechanism: salary-based, task-based, equal-amount, and mixed-method. In this study, I are also interested in exploring RBF treatment effects by the incentive administration types.

4.2 Study design and sampled facilities

The study design is a randomized control trial which I have adapted from the initial RBF design. The MOPH drew the initial RBF design for the purpose of the two household surveys conducted in 2010 and 2012. In early 2010 prior to the start of RBF program, health facilities were randomly assigned into RBF and non-RBF groups in each of the eleven provinces.[29] Randomization process of health facilities within each province is briefly described below.

At the HMIS (Health Management Information System) Department of MOPH in Kabul, the MOPH team stratified health facilities into three strata within each province. The strata were based on the facility type – SCs (Sub Centres), BHCs (Basic Health Centres), and CHCs (Comprehensive Health Centres) in line with the BPHS guideline. The MOPH team prepared a list of matched-pair facilities based on outpatient department (OPD) visits and facility type within each province.

Using the list of matched-pair facilities, the MOPH team went to the respective province and conducted a public lottery for randomization of health facilities. Before conducting the randomization, the team held seminars to orient the NGO managers, and provincial health officers (PHOs) about the introduction of RBF in the eleven provinces.

On the day of randomization, participants of relevant stakeholders such as heads of health facilities, NGO field management team, and the PHO members were invited. At first, names and ID numbers of each matched-pair facilities were announced. The heads of the matched facilities were asked to come over to the stage. For each matched-pair facilities two identically folded pieces of paper, on each of which the name and ID number of each facility had been printed, were shown to the heads of the two matched-pair facilities. After they confirmed the accuracy of the names and ID numbers, the pieces of paper were refolded. The two pieces of paper were put inside a small container, and the lid of the container was capped. After shaking the container and opening its lid, one piece of paper was randomly selected. The name of the selected facility was allocated to the RBF group and its matched-pair was allocated to the non-RBF group.

[29] These provinces, except Kandahar, were selected based on relatively better security situation. Kandahar province was included due to apparently political pressures. Therefore, selection bias seems likely.

In this way, the MOPH team carried out the randomization in a transparent way. At the end of randomization in each province, the results were recorded and the list of randomized facilities was signed by the PHOs, NGO representatives, and by the MOPH team. In total 428 matched-pair (214 non-RBF and 214 RBF) facilities were included in the sample for conducting the two household surveys.[30]

For the purpose of this study, I include a sample of 56 RBF facilities with their matched-pair non-RBF facilities in 2010 (Figure 4.2). This is because I only have data from the 112 facilities over the three years – baseline, mid-line, and end-line surveys – as data from these facilities were collected as part of the NHSPA, not as part of the RBF impact evaluation. The 112 facilities (56 RBF and 56 non-RBF) were surveyed in 2010, so pre-intervention data are available for them. The same sample of 112 facilities were surveyed in 2011and in 2012.

In this way, a panel of health facilities was formed. This can facilitate comparability of changes in health facilities between the control and treatment groups over time. Figure 4.2 shows the 56 RBF facilities selected from the RBF facilities and the 56 non-RBF facilities selected from the non-RBF facilities which were surveyed in 2010, 2011, and 2012. Comparison of findings from the 56 RBF and 56 non-RBF facilities over time yields the RBF treatment effect on the outcomes of interest.

As described in chapter 1, and briefly mentioned under the conceptual framework I are also interested in exploring the RBF treatment effects by the four incentive administration types. To do that I specify RBF facilities and their matched-pair non-RBF facilities by incentive administration types. Of the 56 RBF facilities, 22 implemented salary-based, 19 administered task-based, 8 employed equal-amount, and 4 facilities implemented the mixed-method administration (Figure 4.3).[31]

[30] There were 230 RBF facilities of which, 214 facilities had matched-pair non-RBF facilities. The 16 extra RBF facilities were purposefully included in the program, as either they did not match or they were left over in each of the provinces. The 16 facilities are not included in the household surveys, nor are they included in this study.
[31] Three RBF facilities with their three matched-pair non-RBF facilities were dropped in Balkh province because the RBF facilities in Balkh changed their incentive administration type during project implementation.

Figure 4.2 Study design on RBF overall treatment effects

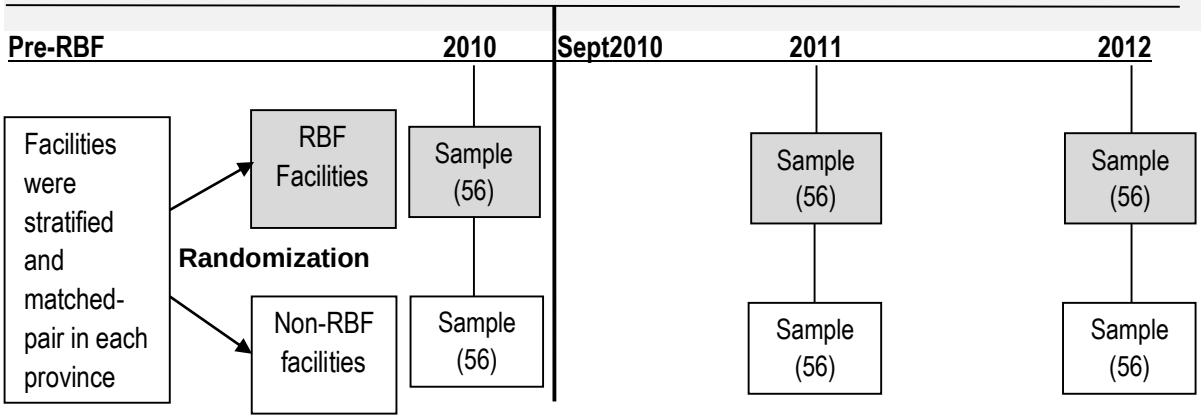
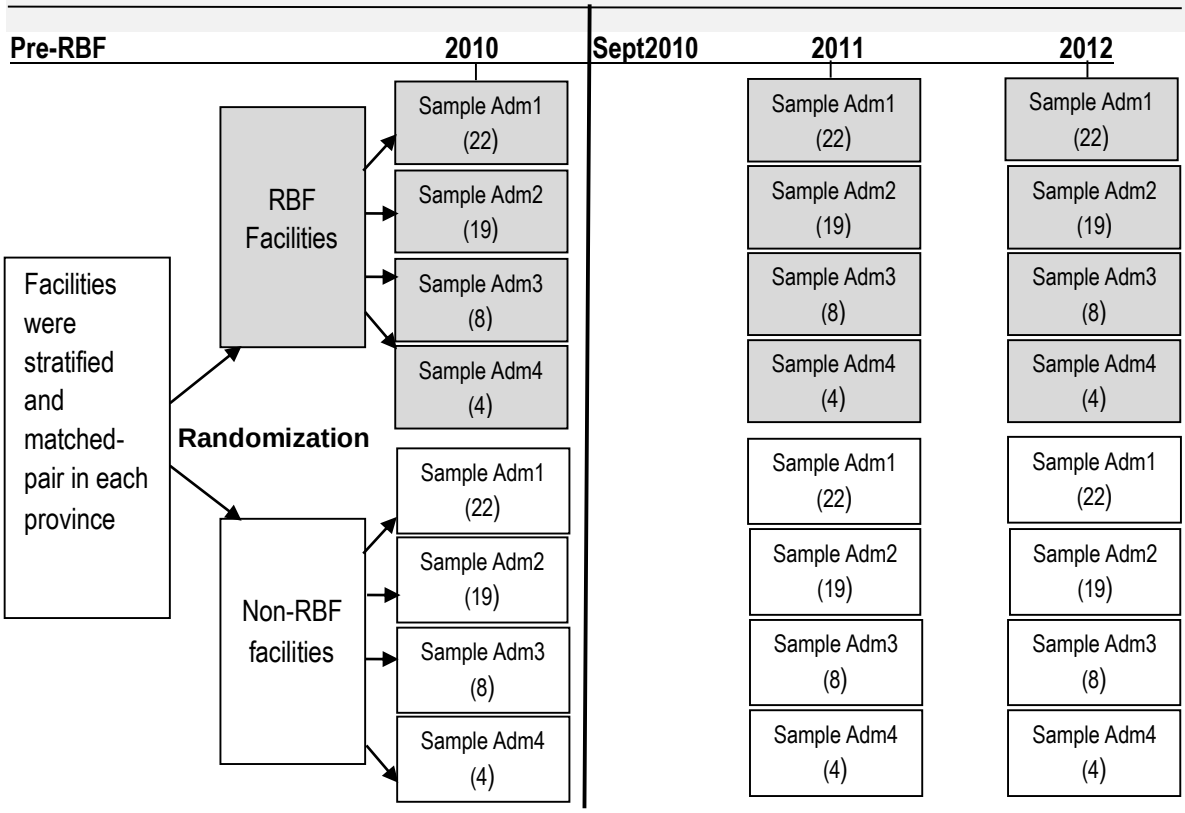


Figure 4.3 Study design on RBF disaggregated treatment effects by the four types of administration



Samples Adm1, Adm2, Adm3 and Adm4 refer to salary-based, task-based, equal-amount and mixed-method administrations respectively

To create control groups for the above four types of RBF administration, I assume that their matched non-RBF facilities would have implemented the same type of administration if they were rewarded the RBF incentives. For example, under the salary-based administration the non-RBF facilities would have distributed incentives among staff according to the salary-based administration if these facilities were given incentives under the RBF program. The same would have been done under the other three admin types. The above assumption is supported by the following argument which highlights the decision power of NGO on selection of incentive administration type for the health facilities.

The NGO had influence and power over the selection of RBF administration type because, except for Badkhshan province, all health facilities in a province were managed by a single NGO. As Table 4.1 shows that in 10 out of the 11 provinces, there was a single NGO managing the health facilities in the province. Table 4.1 shows names of NGOs that managed health facilities (RBF and non-RBF) in the 11 provinces. As can be seen in nearly all cases the NGOs were influential in determining the type of RBF administration. For example, the 12 RBF facilities in Badakhshan (cluster 1) and the 26 RBF facilities in Takhar province were managed by CAF. All of these facilities implemented the equal-amount administration due to the apparent influence of CAF. And yet in Badkhshan we see two types of administration because Badkhshan cluster1 was managed by CAF, and Badkhshan cluster 2 was managed by AKHS.[32] This suggests that under the influence of CAF and AKHS two different incentive admin types for health facilities were administered in Baskhshan.

In addition, nearly in all provinces we see that one type of incentive administration was applied for the respective health facilities. From the management point of view, perhaps it is much easier for a NGO to apply one type of administration in all facilities in the province(s) under the NGO's management power. This further supports the logic that NGOs had control over the selection of administration types for health facilities. Furthermore, in another instance we can see that the 42 RBF facilities in Balkh province, managed by CHA in the first one and half years, employed the salary-based administration. However, due to the take-over of health facilities by AADA in the last year of RBF, the AADA management seemingly influenced RBF facilities to employ the task-based administration, as AADA had already been implementing the task-based administration in Bamyan province.

[32] Badakhshan is a large and mountainous province; therefore, two NGOs operated there.

Table 4.1 Number of RBF and Non-RBF facilities managed by the NGOs in the eleven provinces

NGO	Province	Salary-based	Task-based	Equal-amount	Mixed-method	Total RBF	Total Non-RBF
CAF	Badkhshan(cluster 1)			12		12	10
CAF	Takhar		1	26		27	24
AKHS	Badkhshan (cluster 2)		17			17	12
CHA/AADA	Balkh	42				42	40
AADA	Bamyan		21			21	20
SAF	Jawzjan	10				10	10
AHDS	Kandahar				4	4	4
Merlin	Kunduz	23				23	24
SM-MOPH	Panjshir	8	1			9	8
SM-MOPH	Parwan	29		2		31	29
CSA	Samangan	9	5			14	14
CSA	Sar-e-pul		20			20	20
Total		121	65	40	4	230	214

AADA took over from CHA later and may have influenced RBF facilities to switch to the task-based administration.
The forty two RBF facilities in Balkh province switched to the task-based administration in the last year of RBF.

The above discussion suggests that it is most likely that non-RBF facilities in a province would have implemented the same administration type as their matched-pair RBF facilities. Therefore, it is reasonable that we select the matched-pair non-RBF facilities under each type of incentive administration as the control group for the respective admin type. Of the 56 non-RBF facilities, 22 were selected as control group for salary-based, 19 for task-based, 8 for equal-amount, and 4 facilities as control group for mixed-method administrations. Except for 3 health facilities in Balkh, all others maintained the baseline admin type employed throughout the years. In order to maintain consistency, I dropped the 3 RBF facilities and their matched-pair non-RBF from the disaggregated analysis.

4.3 Survey participants and response rates

In each of the sampled health facilities, data on patient satisfaction were collected from up to ten patients (five patients over 5 years of age, and five patients under 5 years of age). These were selected through systematic random sampling in each round of data collection. For children, their parents or guardians were selected for the exit interviews. Over the three years, around 3000 participants were interviewed; 1033 in 2010, 995 in 2011, and 969 in 2012 (Table 4.2). Table 4.2 shows that response rate in any single year was over 95%.

Table 4.2 Response rates for patient satisfaction and patient – provider interaction

Type of survey	2010	2011	2012	Total
Patient satisfaction	N=1033	N=995	N=969	N=2979
	98%	96%	99%	98%

4.4 Data collection and instruments

Data used in this dissertation come from the health facilities which were surveyed annually under the National Health Service Performance Assessment (NHSPA) program in Afghanistan. Data used in this dissertation were collected from all patients. This point is important when it comes to the interpretation of results. That is, patient satisfaction has been not been stratified by type of services utilized by patients. Therefore, it is not possible to stratify patient satisfaction by incentivized versus non-incentivized services.

For the NHSPA, the Afghan MOPH contracted the Johns Hopkins University and Indian Institute for Health Management Research to conduct the NHSPA survey. [33] The NHSPA was carried out annually in the 34 provinces of Afghanistan. The product of NHSPA was the balanced scorecard (BSC). The BSC is a monitoring and evaluation instrument and was designed through a series of discussions and workshops with the experts from the MOPH, NGOs and other stakeholders involved in the health sector in Afghanistan. The BSC has played a pivotal role in the MOPH's monitoring and evaluation system (Peters et al., 2007; Hansen et al., 2008; Edward et al., 2011). According to Peters et al. (2007), at the design workshops, six domains were identified for incorporating into the BSC. Under these domains, the BSC has 29 key performance indicators (Peters et al., 2007; Edward et al., 2011). Following are the list of the six domains and their indicators respectively.

1. Patient perspectives
    1. Overall patient satisfaction
    2. Patient index
    3. Written health committee activities in community

[33] Ethical approval for the NHSPA was obtained from the Afghan MOPH, and from Johns Hopkins University.

2. Staff perspectives
    4. Health worker satisfaction index
    5. Salary payment current
3. Capacity for service provision (structural inputs)
    6. Equipment functionality index
    7. Drug availability index
    8. Family planning availability index
    9. Laboratory functionality index (CHCs and hospitals)
    10. Staffing index – meeting minimum staff guideline
    11. Provider knowledge score
    12. Staff received training in last year
    13. HMIS use index
    14. Clinical guidelines index
    15. Infrastructure index
    16. Patient record index
    17. Facilities having tuberculosis register
4. Service provision (technical quality)
    18. Patient history and physical examination index
    19. Patient counselling index
    20. Proper sharp disposal
    21. New outpatient visit per month (BHC>750 visits)
    22. Time spent with patient (>9 minutes)
    23. BPHS facilities providing antenatal care
    24. Delivery care according to BPHS
5. Financial systems
    25. Facilities with user fee guidelines
    26. Facilities with exemptions for poor patients
6. MOPH overall vision for the health sector
    27. Females as % of new outpatients
    28. Outpatient-visit concentration index
    29. Patient satisfaction concentration index

Apart from indicators 1, 5, and 22, every indicator was comprised of several related items. To illustrate the point, we report the list of items that were used for calculating indicator 12 (staff received training in last year). The items on which staff received training were: integrated management of childhood illnesses, HIV/AIDS, tuberculosis, malaria, family planning, maternal and new born health, mental health, and disability.

To ensure reliability and validity, the indicators were developed through several discussions with the relevant stakeholders in Afghanistan. According to Peters et al. (2007), the relevant stakeholders assessed an initial list of 340 potential indicators for face validity and reliability. The short list of indicators was then assessed as a group to ensure there was a good balance among those assessing structures, processes, and outputs. After several discussions with stakeholders on indicator definition and selection, a final list of indicators was produced (Peters et al. 2007).

To further ensure face validity and reliability of survey forms, all items, instructions, skip patterns, and codes in the survey forms were checked both for Dari and Pashto (local languages) forms to be matching with those of the English version. The survey forms were developed by experts from The Johns Hopkins University, Indian Institute for Health Management Research, and Monitoring Department of the MOPH in Kabul. The survey forms were translated into Dari and Pashto. The Dari and Pashto version of the survey forms were back-translated into English by a different translator who was neither involved in translating the original English forms into Dari or Pashto, nor did he/she have any exposure or access to the original English survey forms. The experts re-checked the translation to ensure content validity and reliability of the survey forms. In the field, either Dari or Pashto version of the survey forms were used.

For the purpose of the BSC, a sample of up to twenty five BPHS facilities was randomly selected for each province. To avoid possible selection bias during sampling, the sampling was carried out by the NHSPA program manager in Kabul. Selection bias occurs when subjects in a study are not randomly selected from the population under study. Unless corrected by the use of sophisticated statistical modelling, selection bias can result in either overestimation or underestimation of a treatment effect (Heckman 1997). For the NHSPA, using an updated list of public health facilities which was obtained from the health management information system (HMIS) department of MOPH, the NHSPA program

manager carried out the sampling. The program manager used the Microsoft Excel program, and generated random numbers for the sampling. The sampling method was stratified simple random sampling. Using the MOPH standard definitions and categories of health facilities, within each province the BPHS facilities were stratified into their three types: CHCs, BHCs, and SCs. Up to 5 SCs, 15 BHCs, and 5 CHCs were randomly selected for the NHSPA in each province.[34]

Field work for the NHSPA: To provide a complete picture of the assessment of public health facilities, data related to structures, processes, and output/outcomes were collected from health facilities in each province. Using the list of sampled facilities, the surveyors collected data through reviewing records/registrations, observing processes and resources (e.g., providers interaction with patients, availability of medicine and supplies, etc.), and by interviewing health providers and patients in health facilities.

Each participant (the patient or caregiver who was selected) was asked for informed consent. For children, the parent or caregiver was asked for the informed consent. To get informed consent, the surveyors explained the purpose and procedure of data collection, and the rights of participants to decide about his/her participation or withdrawal from the survey.

After obtaining informed consent, the surveyor observed the patient - provider interaction (e.g., taking history of patient, physical examination of patient, etc.), and filled in the survey forms. Except for the observation, the surveyor was trained not to ask any question, nor talk to the health provider or patient during the patient – provider interaction. An exit interview with the same patient (or caregiver) was conducted after the patient received the prescribed medicine and was about to leave the facility. To minimize the likelihood of bias, a different interviewer conducted the exit interview.

To measure levels of patient satisfaction a questionnaire with four-point Likert-scale was used (very satisfied, satisfied, dissatisfied, very dissatisfied). Likert type of questions with different scales are commonly used in customer satisfaction surveys in the health sector (Peterson & Wilson 1992). How a respondent distinguishes between very satisfied and satisfied or between dissatisfied and very dissatisfied is a matter of subjectivity, depending on

[34] The product of NHSPA is the BSC.

the respondent's interpretation of the scale (Cronin & Taylor., 1992). To ensure, the respondents understand the concept of four-point Likert-scale, a laminated "naan-scale" card was used. One, two, three and four "naan", which is the Dari word (or dodai in Pashto) for bread, were assigned to the "naan-scale" card. It was explained to the respondent that one "naan" corresponds to "very dissatisfied", two "naan" means "dissatisfied", three "naan" stands for "satisfied" and four "naan" displays "very satisfied" for any items in the list. The surveyor read each item to the respondent, and asked him/her to point out to the number of "naan" he/she would assign to the item. Then the surveyor circled the respective answer in the survey form.

The interviews were conducted in either Dari or Pashto at the discretion of respondents. The interviewers were fluent in speaking both languages. They asked respondents about their preference to have the interview in Pashto or Dari. The two languages are commonly spoken in nearly all over Afghanistan, and it is rarely the case that a person cannot speak either of the languages. The option for patients to choose their preferred language during interview can ensure that respondents feel comfortable, and understand the questions.

Monitoring of survey teams: The survey teams were monitored through two mechanisms: (i) active monitoring, and (ii) post monitoring. Active monitoring refers to the concurrent collection of data from a certain number of sampled facilities by a survey team and a field monitor.[35] Post monitoring refers to the reassessment of a certain number of sampled facilities by the field monitor after several weeks following the assessment of sampled facilities by the survey teams.

During the active monitoring, a field monitor accompanied a survey team to up to five sampled facilities and collected data concurrently. Following a random selection of up to five facilities by the NHSPA program manager in Kabul, the field monitor prepared his itinerary and departed for active monitoring. In each facility, the field monitor collected data from two patients who were selected by the survey team. One of the two patients was under 5 years old, and the other was over 5 years of age.

[35] Individuals with medical background had been recruited and trained by the Johns Hopkins University team in Kabul to work as field monitors.

Data were collected in two ways: (i) direct observations, and (ii) exit interviews. The interaction between the health provider and patient (caregiver) was observed by the surveyor and the field monitor. The surveyor and field monitor did not consult with each other during this process. Throughout the observations, the field monitor and surveyor filled in the survey forms. The same patient was interviewed for patient satisfaction measures by a second surveyor and the field monitor.

The field monitor compared the responses in his survey forms with those obtained by the surveyors at the end of survey. If the level of agreement between the records of the two survey forms was less than 50%, the field monitor asked the survey team to stay in the facility and recollect data – either on the same or the following day. The field monitor also provided technical assistance and feedback to the survey team.

For the post monitoring, a field monitor reassessed up to five facilities out of the sampled facilities which had been surveyed several weeks ago. Unlike the active monitoring, those characteristics of facilities were assessed that were less likely to change over a period of several weeks. For example, variables such as number and type of staff, availability of essential drugs and supplies were assessed. The data collected through the active and post monitoring missions were analysed for inter – rater reliability testing in Kabul.

Inter - rater reliability tests: Inter-rater reliability testing refers to the degree of agreement between the data collected from a facility by a survey team and the data collected from the same facility by a field monitor. When the agreement between data collectors is high, the result of inter-rater reliability testing also tends to be high. This testing provides information about the accuracy of surveyors work in health facilities. For the inter-rater reliability I provide percentage agreement and Kappa statistics below.

Percentage agreement refers to the extent of concordance between the responses obtained by a survey team and a field monitor from the same facility. Kappa statistics also refers to the same concept, but, it eliminates the occurrence of agreement by chance. This makes Kappa statistics a better measure of inter-rater reliability than the percentage agreement (Gordis. Leon, 2009).

The results obtained from inter-rater reliability testing were used to improve quality of data. Percentage agreement between the data collected by the surveyors and the field monitor, and Kappa statistics were calculated to determine the level of inter-rater reliability. A fifty percentage agreement which shows a fair level was considered to be the minimum level of agreement for inter-rater reliability testing.

Here I provide the results of inter-rater reliability testing only for the 11 RBF provinces. For 2010 I can only report the percentage agreement provided by The Johns Hopkins University team in Afghanistan. As I did not have access to the monitoring database of 2010, I cannot provide Kappa statistics for the year. For 2011 and 2012, I present the percentage agreement and Kappa statistics through analysing data for the inter-rater reliability testing. At first, I calculated the percentage agreement and Kappa statistics at facility level, and then I computed average of percentage agreement and average of Kappa statistics to obtain these measures at province level. Table 4.3 shows percentage agreement and Kappa statistics for active monitoring and post monitoring at province level over 3 years.

The results show that percentage agreement and Kappa statistics were 50% and over for the 11 provinces in 2010. Badakhshan was the only province with 50% percentage agreement which is a fair agreement in terms of inter-rater reliability. This seems to be a reason that more monitoring visits to health facilities were conducted in the following two years.

Table 4.3 Percentage agreement and the Kappa statistics for post and active monitoring at province level

Province	Post monitoring					Active monitoring				
	2010	2011		2012		2010	2011		2012	
	%	%	kappa	%	kappa	%	%	kappa	%	kappa
Badakhshan				80	70	50			95	90
Bamyan	74	75	70				86	80		
Balkh	70									
Jawzjan	73			78	70				90	80
Kandahar	59			80	70	78			86	80
Kunduz	66			77	70		91	90	84	80
Samangan	70	73	60	73	60				97	90
Sar-e-pul	79	63	50	79	70		97	90	92	80
Panjshir	67			76	60		92	90	86	80
Parwan	74			73	70		79	70	88	80
Takhar	65			79	70		90	90	95	90

In 2011 and 2012 the degree of agreement between the records of survey teams and field monitors increased. The results in Table 4.3 show that higher percentage agreement and Kappa statistics were obtained for the active monitoring, followed by post monitoring. This may be attributed to more technical assistance as the field monitors provided more direct feedback and guidance to surveyors during active monitoring. The results from the post monitoring also show higher degree of agreement in terms of inter-rater reliability. That is, for post monitoring Kappa statistics are 60% and over which indicate substantial agreement, and for active monitoring Kappa statistics are 75% and over which show excellent agreement (Gordis. Leon, 2009, p. 105).

The fair to moderate agreement in 2010, and the substantial to perfect agreement in 2011 and 2012 suggest that quality of data improved over time. In 2010, even if the degree of agreement was lower than the other two years, quality of data was maintained as suggested by the presence of a percentage agreement of nearly 60% and over.

4.5 Quality of data

To ensure collection of quality data, various measures were employed by The Johns Hopkins University team in Afghanistan. These measures included recruiting competent staff, ensuring validity and reliability of survey forms, supervising and monitoring surveyors, and editing the data collected from the field.

Recruiting competent staff: The Johns Hopkins University team in Kabul hired and trained surveyors before deploying to the sites. The majority of potential applicants had medical background. A competency test was set up, and the applicants who passed the competency test at the end of the training sessions were recruited for data collection. Based on their level of competency the successful applicants were hired as survey supervisors, field editors, and surveyors. In each survey team there was one supervisor, two surveyors, and one field editor. A survey team was deployed for each province and they collected data from the sampled facilities (sampling of health facilities was discussed earlier).

Reliability and validity of survey forms: Reliability of an instrument refers to its consistency in getting the information it intends to get. Reliability of an instrument can be seriously impaired by poorly worded and/or unclear questions or instructions, which results

in misunderstanding and misinterpretation of the questions asked, and the information obtained (Fink 2003). An unreliable instrument can also be invalid. Validity refers to the accuracy of an instrument in measuring what it is supposed to measure (Fink 2003). Validity and reliability was also taken into serious account for ensuring quality of data. For this, experts from The Johns Hopkins University, Indian Institute for Health Management Research, and the Afghan MOPH worked on the survey forms. They held several sessions in Kabul where they discussed the relevance of each item for respective domain, and developed the survey forms through a consultative process. The experts from Afghan MOPH were from the Monitoring Department and GCMU who shared theirs insights and experiences and those of the implementing partners (e.g., NGOs) and local authorities such as PHOs during development of the survey forms. The experts from The Johns Hopkins University and Indian Institute for Health Management Research had sufficient knowledge of public health system, and context of Afghan society. In this way, it was ensured that the items in the survey forms were relevant for the public health sector in Afghanistan.

The survey forms were field tested throughout the development process. They were field tested in different types of health facilities and in several districts to ensure the reliability and consistency of the forms. The results obtained from the field tests were discussed with the experts, and the items were revised accordingly. Any ambiguity in the interpretation of items was discussed and the items were revised accordingly. For example, a double-barrel question can confuse a respondent or data collector;[36] therefore, such questions were broken down in to simple and easily-comprehensible items. The concept of double-barrel items and mixed wording is particularly important when it comes to patient satisfaction. The mixed wording may reduce the potential "yes" and "no" responses from respondents; but, it can also confuse them. This is particularly the case, when the items are negatively worded as the respondents may require more time to understand the questions, and this can lead the respondents to make more comprehension mistakes (Babakus & Boller, 1992). Therefore, efforts were made to avoid negatively worded items with respect to patient satisfaction.

[36] An example of a double-barrel question is "How satisfied are you with cleanliness of and accessibility to facility?" This question can be broken down into two questions: "How satisfied are you with cleanliness of facility" and "How satisfied are you with accessibility to facility?"

During 2010, the survey forms including the items on patient satisfaction went through a review and revision process. Most of the questions were reworded from 'statement' to 'question' forms. To test the reliability of the instrument, I have conducted an internal consistency test of the twelve items on patient satisfaction. The result shows an alpha coefficient of 0.78 in 2010, and an alpha coefficient of 0.80 in 2011. This suggests an acceptable level of internal consistency between the twelve items. Conventionally, an alpha coefficient of more than 0.80 is considered to be a good result for reliability of an instrument (Carmine & Zeller, 1979, p. 51).

I also conducted construct validity assessment of the instrument (Table 4.4). To test the instrument for construct validity, the SERVQUAL framework, developed by (Parasuraman et al., 1988, 1991) has been used as a guide. According to SERVQUAL scale, covers mainly five dimensions: tangibles, reliability, responsiveness, assurance, and empathy. These dimensions have served as a guide in some studies on patient satisfaction in developing countries; however, the number of dimensions have not been consistent in these studies. Carman (1990) suggests that based on the context and situation, dimensions of SERVQUAL can be modified or deleted, or new dimensions can be added. Using the three years' dataset, I have employed the principal component analysis to identify the number of components (dimensions), and found the following results. A principal component is defined as a linear combination of optimally-weighted observed variables (Harman, 1976).

Table 4.4 The highest factor loadings on the respective components related to patient satisfaction

Variable	Empathy	Responsive	Assurance	Tangible	Expenses
Patient satisfaction with					
Traveling from home to the facility				0.82	
Cleanliness of the facility				0.48	
Respectfulness of the health provider			0.51		
Privacy in the facility			0.74		
The way health provider explains illness	0.69				
The way health provider explains treatment	0.66				
Cost of visiting the facility					0.59
Convenience of getting the prescribed drugs					0.63
Sufficiency of the visit time		0.38			
Adequacy of open hours of the facility		0.65			
Waiting time in the facility		0.62			

The eleven items were associated at varying degrees of factor loadings on the five components; however, I report the highest factor loadings of the items on the respective component to bring clarity in the results

I used an eigenvalue of 0.9 during principal component analysis to retain sufficient number of items in the coefficient matrix. After rotating the coefficient matrix I obtained the five components on which the eleven items loaded (Table 4.4). The first component mainly represents items related to 'empathy', the second component primarily contains items reflecting 'responsive', the third component deals with 'assurance', the fourth component subjectively refers to 'tangible', and the fifth component mainly covers the 'expenses' dimension of patient satisfaction. The above five dimensions resemble the ones proposed by the SERVQUAL. This suggests that the above items can measure what they are supposed to.

4.6 Estimation model

A regression adjusted Difference-In-Differences (DID) estimator model is used to investigate the RBF treatment effect. The DID estimator calculates the effect of a treatment by comparing the mean change in the outcome over time between the treatment and control groups. The DID estimator is based on the identification assumption that in the absence of treatment the average change in the outcome for the treated and control groups would have followed parallel paths over time (Abadie, 2005). The identification assumption can be valid when the pre-treatment observed characteristics between treatment and control groups have a balanced distribution.

The DID estimator removes biases due to time invariant characteristics specific to treatment and control groups (Meyer, 1995), as the estimator links treatment effect to change over baseline for each group. For example, in using panel data from health facilities when the same facilities within each group of RBF and non-RBF is observed in 2010 and then in 2011 and 2012, the average gain in the outcome in non-RBF group is subtracted from the average gain in the outcome in RBF group. This removes biases in 2011 and 2012 between RBF and non-RBF groups which could be as a result of permanent differences between the two groups, as well as biases emanating from comparisons over time in RBF and non-RBF groups that could be due to trends. Following is a simple demonstration of the DID model.

$$DID = [(\text{mean}^{\text{RBF}}{}_{2012}) - (\text{meant}^{\text{RBF}}{}_{2010})] - [(\text{mean}^{\text{non-RBF}}{}_{2012}) - (\text{mean}^{\text{non-RBF}}{}_{2010})]$$

The regression model below detects effect size over the first and second years on the outcome of interest.

$$Y_{ijt} = \beta_0 + \beta_1 RBF_{jt} + \beta_2 yr1_{jt} + \beta_3 yr2_{jt} + \beta_4 RBF * yr1_{jt} + \beta_5 RBF * yr2_{jt} + \Sigma_{k=1}^{K} \beta_j X_{ijt} + \varepsilon_{ijt}$$

Where

Y_{ijt}: The outcome variable for respondent *i* facility *j* at time *t*

RBF_{jt}: Dummy variable for RBF in facility *j* at time *t*

$yr1_{jt}, yr2_{jt}$: Dummy variables for year 2011 and 2012 (2010 is the reference)

$RBF * yr1_{jt}, RBF * yr2_{jt}$: Interaction terms between RBF and year 2011, and between RBF and year 2012

X_{ijt}: A vector of control variables

β_4, β_5: Treatment effects associated with RBF in 2011 and 2012 respectively

4.7 Choice of regression models

The choice of regression model depends on how the outcome variables are measured or constructed. For the Likert scale responses of patient satisfaction, I use ordered logit model and provide marginal probability effect on each of the 12 items. For the outcome variables with binary response, I use the linear probability model (LPM) which is also used for the continuous outcome variables, and the model is then called the ordinary least square (OLS).

The ordered logit model is often used when the outcome variable has an ordinal scale such as "very satisfied", "satisfied", "dissatisfied", and "very dissatisfied" in the case of my study. The ordered logit model allows for "rank order" between the scales of responses. In my study, using the DID approach with the ordered logit model, I provided marginal probability effect on the four scales of responses. The marginal probability effect measures RBF treatment effect on each items of patient satisfaction at the four scales of responses. Here I

provide an example on marginal probability effect on overall patient satisfaction (Table 4.5). Table 4.5 shows that in 2011, marginal probability effect on overall patient satisfaction was – 0.005 at the "very satisfied" scale. On the other three scales, marginal probability effect was with positive sign: it was 0.004 at the "satisfied" scale, 0.001 at the "dissatisfied" scale, and 0.000 at the "very dissatisfied" scale of responses. In 2012, marginal probability effect was 0.095 at the "very satisfied" scale; but was with negative sign for the other three scales: it was – 0.076 at the "satisfied" scale, – 0.017 at the "dissatisfied" scale, and – 0.002 at the "very dissatisfied" scale of responses. The same pattern was observed between the sign of "very satisfied", and the sign of other three scales of responses for all items of patient satisfaction. Sum of the marginal probability effects across the four scales is zero in 2011 and 2012. This means that an increase in "over satisfied" scale is accompanied by a combined decrease in the marginal probability of the other three scales, and vice versa.

Table 4.5 Marginal probability effect of RBF on overall patient satisfaction

Treatment Effect	Very satisfied	Satisfied	Dissatisfied	Very dissatisfied	pseudo R^2	N
in 2011	-0.005 (0.066)	0.004 (0.053)	0.001 (0.011)	0.000 (0.002)	0.012	2958
in 2012	0.095 (0.074)	-0.076 (0.059)	-0.017 (0.013)	-0.002 (0.002)	0.012	2958

Ordered logit model was used. Standard errors are provided in parenthesis.

From the findings in Table 4.5, it is clear that it becomes cumbersome to interpret RBF treatment effect on patient satisfaction at each scale of responses. Therefore, a meaningful way of estimating RBF treatment effect is to provide marginal probability effect at "very satisfied" scale because any increase in this scale is associated with decreases in the other three scales combined, and vice versa. Another way to estimate RBF treatment effect is to use binary response for each of the patient satisfaction items, as this allows for relative degree of freedom between the two responses (yes, no). For this, I recode the four-point Likert scale responses to binary variable and use the LPM. In chapter 6, results from both LPM and ordered logit model are provided to check consistency in the results between the two models. The LPM not only allows for binary response, but it can also be used for continuous variables.

4.8 Variables

Outcome variables in chapter 6 are the twelve items on patient satisfaction, which include one item on overall satisfaction. Marginal probability effect at the "very satisfied" scale has been produced for each of the 12 items. The four-point scale responses have been recoded as binary (very satisfied=1, and other three responses=0) for the use of LPM. Also aggregate indices have been computed and used for the descriptive statistics in chapter 6, and as the outcome variables in chapter 7. The choice in chapters 6 and 7 that which variables to be included in the respective aggregate index is determined according to the result obtained for the construct validity (Table 4.4). The results on construct validity show that there are five components: empathy, assurance, tangible, responsive, and expenses. The aggregate indices are weighted average score of the items included in the index. A composite index which is the weighted average score of all items, except the overall satisfaction, is also computed and used in the analysis.

In chapter 8, the outcome variables are the three aggregate indices related to health provider performance; the three aggregate indices related to availability of medicine, vaccines, and equipment; and the two aggregate indices related to physical condition, and cleanliness of health facilities. The choice based upon which related items were included in the respective index is decided by the relevance of the items to the index. For example, all categories of essential drugs are included in the medicine index. The aggregate indices have been computed as the weighted average score of the items included in the index. In addition, three composite indices (composite index for health provider performance, availability of the 3 types of resources, and physical appearance of health facilities) have been computed as the weighted average score of related indices (or all related items).

The control variables included in the model are thought to be potentially exogenous of RBF treatment effect on the outcome of interest. In chapters 6, and 7 where the outcome variables are items of patient satisfaction, the control variables included in the model are: ability of respondents to read and write, age of patients, sex of patients, socioeconomic status of patients, and distance of a patient house from the health facility. Patient age and distance of a patient house from the health facility were continuous, and other control variables were dummy variables. In chapter 8, where the outcome variables are health provider performance, the control variables are the same used in chapters 6 and 7. However, when the outcome

variables are availability of medicine, vaccines, and equipment, and physical appearance of health facilities, the control variables are: type of the health facility (whether the facility is a comprehensive health centre), and distance of the health facility from the provincial capital.

4.9 Clustering effects and adjustment of standard errors

Consequence of clustering effects in using panel data has attracted a lot of attention in empirical studies. In our study, clustering of data within health facilities can happen due to the fact that patients interviewed share similar characteristics because they live in the catchment areas of respective facility. This results in clustered data and intra-cluster correction which is accompanied by lesser variability in the information provided by patients.

This problem leads to biased estimation of standard errors and subsequent statistical tests. Therefore, adjusting variation between facilities which refers to adjusting standard errors at health facility level is needed to account for the intra-cluster correlation. In my study I adjust standard errors at health facility level by using the bootstrap command in STATA, and provide clustered standard errors (Gustavo Sanchez, June 2013).

Chapter 5: Comparison of baseline

This chapter aims to examine whether there was a balanced distribution of characteristics between RBF and non-RBF facilities in 2010. In principle randomization is expected to provide balanced distribution of covariates, but in practice it may not always lead to balanced distribution so it is always preferred to check. In this chapter I check whether such a balanced distribution existed during baseline when health facilities were randomly allocated to RBF and non-RBF groups. The chapter is comprised of three sections: section one compares key characteristics of health facilities, section two compares background information of patients, and section three compares key attributes of health workers between RBF and non-RBF facilities in 2010.

5.1 Health facility attributes

In order to compare characteristics of health facilities in the baseline period I utilize data from the facility assessment datasets developed for the NHSPA program. I provided average monthly use of target services between RBF and non-RBF facilities in 2010 (Table 5.1). The target services are antenatal care (ANC) visits which are 4 rounds, postnatal care (PNC) visits which are 2 rounds, Diphtheria-Pertussis-Tetanus 3rd round (DPT3) vaccination, institutional deliveries, and detection of tuberculosis cases (Table 5.1).

Table 5.1 shows that there was no significant difference between the average monthly use of target services between RBF and non-RBF groups in 2010.[37] For example, the monthly mean for ANC1 visits was 36 cases in RBF and 35 cases in non-RBF facilities. Likewise, the monthly mean for PNC1 visits was 16 cases in RBF and 16 cases in non-RBF. For each type of target services, the difference was not larger than 1 case between the two groups of health facilities.

The 56 RBF and 56 non-RBF facilities were each composed of 34 BHCs, 14 CHCs, and 8 SCs. Using data from these facilities, I compared the two groups by managing agencies, by distance from the provincial capital, and by patient loads in 2010 (Table 5.2). Table 5.2 shows that the percentage of RBF versus non-RBF facilities managed by organizations were the same. For instance, 7% of health facilities in RBF and the same percentage of health

[37] They are statistically confirmed by findings from the t-test and p-values

facilities in non-RBF were managed by the MOPH. The proportion of health facilities managed jointly by the MOPH and NGOs was 28% in RBF, so was in non-RBF. Percentage of health facilities managed by NGOs alone formed the biggest proportion: it was 65% in RBF and 65% in non-RBF facilities. No difference was seen between the two groups of facilities in terms of managing agencies in 2010.

Percentages of health facilities by distance (in kilometres) from the provincial capital were similar between the two groups of facilities.[38] Twenty four percent of RBF and 28% of non-RBF facilities were located less than 19 kilometres from the provincial capital. Similarly, 19% of RBF and 12% of non-RBF facilities were located between 20 to 39 kilometres from the provincial capital. Percentages of RBF and non-RBF facilities located between 40 to 59 kilometres away from the provincial capital were 18% and 10% respectively. Thirty eight percent of RBF and 49% of non-RBF facilities were located over 60 kilometres from the provincial capital. The existence of a large p value ($p=0.43$) confirms the insignificant difference between the two groups of facilities.

Distance of health facilities was also estimated based on travel time (in hours) from the provincial capital. Percentages of RBF and non-RBF facilities located less than half an hour from the provincial capital were 12% and 16% respectively. Twenty two percent of RBF and 14% of non-RBF facilities were located between half an hour and an hour from the provincial capital. Percentage of RBF and non-RBF facilities located between an hour and an hour and half were 14% and 7% respectively. Despite a difference of 7%, there is no statistically significant difference between the RBF and non-RBF facilities. Nine per cent of RBF and 12% of non-RBF facilities were located between an hour and half and two hours from the provincial capital. Percentages of RBF and non-RBF facilities located more than two hours from the provincial capital was 43% and 51% respectively. Here too, we see no significant difference between the two groups of health facilities.

In Table 5.2 I also compared RBF and non-RBF facilities in terms of patient load – an indicator of service utilization. For this, I computed mean patient visits in the last six months per health facility.[39] I computed mean patient visits for male and female patients as well as

[38] Data on distance of a health facility were collected through interviews with head of the health facility
[39] Data on patient loads were collected from each facility's registers

for under five year old patients. The mean visits for male patients was 2760 cases in RBF and 2562 cases in non-RBF facilities. The mean visits for females was 5424 cases in RBF and 4831 cases in non-RBF facilities. The mean visits for males under 5 years was 1291 cases in RBF and 1180 cases in non-RBF facilities. Similarly, the mean visits for females under 5 years was 1235 cases in RBF and 1101 cases in non-RBF facilities in 2010. The p values for all the four categories indicated insignificant differences between RBF and non-RBF facilities.

5.2 Patient characteristics

In order to compare characteristics of patients between RBF and non-RBF facilities, I looked into the patient demographic data, and socioeconomic status between the two groups of health facilities in 2010 (Table 5.3). Here I compare patient sex, education level, socioeconomic status, and patient age between the two groups of facilities. In each sampled facility, patients or caregivers were asked about the age and sex of patients, about the number of years of official education, and about the ownership of common household assets.

For this analysis, I created five levels of official education, and three levels of socioeconomic status for patients. The five levels of official education were: 'no schooling', 'primary school' which included one to six years of official education, 'intermediate school' which covered seven to nine years of official education, 'high school' which included ten to twelve years of official education, and 'higher education' which referred to over twelve years of official education. These levels of official education were defined according to the prevailing practice in Afghanistan.

For the patient socioeconomic status, I categorized households based on the number of assets they possessed. The households having up to five assets were categorized as "very poor", the households possessing up to nine assets were classified as "poor", and the household owning more than nine assets were named "not poor". The purpose for this categorization was to classify the households into three socioeconomic types in order to compare them between RBF and non-RBF facilities in the baseline (Table 5.3).

Table 5.3 shows that characteristics of patients were not statistically different between RBF and non-RBF facilities in 2010. There were 45% male patients in RBF and 44% male patients in non-RBF facilities. Similarly, there were 54% female patients in RBF and 55% female patients in non-RBF facilities. Percentage of patients by education level shows that there were 84% patients (or caregivers) without schooling in RBF, and 86% patients (or caregivers) without schooling in non-RBF facilities in 2010. Percentage of patients (or caregivers) who completed primary school was 7% in RBF and over 7% in non-RBF facilities. Similarly, percentage of patients (or caregivers) who completed intermediate school were 5% in RBF and 4% in non-RBF facilities. Percentage of patients (or caregivers) who completed high school was 3% in RBF and over 2% in non-RBF facilities, and percentage of patients with higher education level was less than 1% in RBF and less than 1% in non-RBF facilities. No substantial differences were seen between RBF and non-RBF facilities.

Percentage of patients by socioeconomic status shows that in 2010, 89% of patients attending RBF facilities were very poor, and 91% of patients visiting non-RBF facilities were very poor. Similarly, 11% of patients visiting RBF facilities were poor, and 9% of patients attending non-RBF facilities were poor. With regard to not poor patients, less than 0.5% of patients in RBF and less than 1% of patients in non-RBF facilities were in this category. There was no significant difference between RBF and non-RBF in terms socioeconomic status of patients. Table 5.3 also shows that there was no significant difference in the mean age of patients between RBF and non-RBF facilities. The mean age for male patients, for example, was 16 years old in RBF and was slightly over 16 years old in non-RBF facilities. Similarly, the mean age for female patients was 17 years old in RBF and was slightly over 18 years old in non-RBF facilities.

5.3 Health worker attributes

Health workers are essential for the delivery of health services as the health workers form the locomotive of any health interventions (Henderson & Tulloch, 2008; Luboga et al., 2011). Not only their presence in terms of number and type is important; but, their performance and the way they are managed and supported is crucial. In this subsection, I therefore compare characteristics related to the number and types of health workers, and the ways they were supported technically in 2010.

Background information of health workers and technical supports provided to health workers are compared between RBF and non-RBF groups in Table 5.4. Under the background information, sex of health workers, position of health workers, work experience of health workers, interaction of health workers with their supervisors, and training of health workers will be compared between the two groups of facilities. Table 4.4 shows that there were more male health workers than female health workers in health facilities in 2010. There was 60% male health workers and 40% female health workers in RBF facilities. Similar percentages were seen in non-RBF facilities – it was 61% for male and 39% for female health workers. No visible difference was seen between the two groups.

The composition of health facilities in terms of various types of health workers was similar between RBF and non-RBF facilities. Eighteen percentage of health workers in RBF and 17% of health workers in non-RBF facilities were medical doctors. Nurses constituted a similar portion of health workers in health facilities. Twenty percentage of health workers in RBF and 17% of health workers in non-RBF facilities were nurses. Midwives, on the other hand, formed exactly the same proportions of health workers in both groups of facilities. In both RBF and non-RBF facilities, 24% of health workers were midwives. The largest proportion of health workers concerned the vaccinators – this category of staff formed 30% of health workers in both RBF and non-RBF facilities. The smaller proportions of health workers relate to assistant doctors, CHS (Community Health Supervisors), and pharmacists. With respect to these categories of health workers, here too there was not much difference between RBF and non-RBF facilities. In all, no significant difference in terms of number and types of health workers was seen between the two groups of health facilities.

The composition of health facilities with regard to health workers by work experience was similar between the two groups of health facilities in the baseline. The largest proportion of health workers by work experience was between one year and five years, and the lowest was less than one year of experience. The percentage of health workers having work experience of less than one year was 9% in RBF and 11% in non-RBF, and the percentage of health workers having work experience of between one year and five years was 47% in RBF and 47% in non-RBF facilities. Percentage of health workers having work experience of between six to ten years was 20% in RBF and 21% in non-RBF groups. No significant difference existed between the two groups of facilities with regard to having experienced health workers.

Another important aspect for comparison was the interaction of health workers with a supervisor recently. Here too, similarities between RBF and non-RBF facilities are evidenced. Percentage of health workers who interacted with a supervisor in the past 30 days was 84% in RBF and 83% in non-RBF facilities. Likewise, similar percentages were seen among the health workers who interacted with a supervisor in the past 31 to 90 days – the percentage was 12% in RBF and 9% in non-RBF facilities. Similar proportions, though very small, were observed among the health workers who had interacted with a supervisor in the past three to six months, and more than six months. A similar situation was seen among the health workers who did not have any interaction with a supervisor in the past. The existence of a large p value (p=0.38) confirms the insignificant differences between RBF and non-RBF facilities in terms of health workers who interacted with their supervisors.

A relevant aspect for comparison was the job related training health workers received recently. Like the above, similarities can be seen between RBF and non-RBF facilities. Percentage of health workers who received job related training in the past 7 days was 33% in RBF and 30% in non-RBF facilities. Likewise, percentage of health workers who did not receive any training was similar between the two groups – it was 20% in RBF and 25% in non-RBF facilities. Similarly, 22% of health workers in RBF and 25% of health workers in non-RBF facilities received job related training in past 8 to 14 days. Percentage of health workers who received job related training in the past 15 to 21 days was the same between RBF and non-RBF facilities (12% in each group). Percentage of health workers who received job related training over 21 days ago was 11% in RBF and 7% in non-RBF facilities. The existence of a large p value (p=0.60) indicates no significant difference between the two groups of facilities in terms of providing job related training to health workers.

5.4 Conclusion

The evidence seen in the three sections indicates that overall there were no significant differences between the two groups of facilities in the baseline. This suggests that randomization of health facilities into RBF and non-RBF groups achieved the intended objective of balanced distribution of attributes related to health facilities, patients, and health workers. This suggests that the likelihood of selection bias in terms of various confounding factors was minimized in the baseline.

Table 5.1 Monthly mean of target services between the RBF and non-RBF groups in 2010

Health services	RBF N=56 Mean	Non-RBF N=56 Mean	Difference	P-value
ANC1 visits	35.8	35.2	0.6	0.89
ANC2 visits	14.6	15.8	-1.2	0.59
ANC3 visits	8.3	9.0	-0.7	0.65
ANC4 visits	4.9	5.2	-0.3	0.77
PNC1 visits	16.5	16.4	0.1	0.99
PNC2 visits	6.2	7.1	-0.9	0.58
Institutional deliveries	8.4	9.4	-1	0.63
DPT3 coverage	48.9	48.4	0.5	0.94
TB case detection	0.43	0.37	0.06	0.71

Table 5.2 Characteristics of health facilities in RBF and non-RBF groups in 2010

Variable	RBF N=56	Non-RBF N=56	Difference	P value
Percentage of facilities by managing agency	%	%		
MOPH	7	7	0	
MOPH and NGOs	28	28	0	0.99
NGOs	65	65	0	
Percentage of health facilities by distance from the provincial capital (in kilometres)				
< 19	24	28	-4	
20 to 39	19	12	7	
40 to 59	18	10	8	0.43
> 60	38	49	-11	
Percentage of health facilities by travel time from the provincial capital (in hours)				
< 0.5	12	16	-4	
0.5 to 1	22	14	8	
1 to 1.5	14	7	7	0.49
1.5 to 2	9	12	-3	
> 2	43	51	-8	
Patient visits in six months per facility	Mean	Mean		
Male patient visits	2760	2562	198	0.77
Female patient visits	5424	4831	593	0.20
Patient visits (male under 5 years)	1291	1180	111	0.64
Patient visits (female under 5 years)	1235	1101	134	0.55

Chi-square test and t-test were used for the categorical and continuous variables respectively

Table 5.3 Characteristics of patients in RBF and non-RBF health facilities, baseline – 2010

Variable	RBF facilities		Non-RBF facilities		
	%	N	%	N	P value
Percentage of patients by sex					
Male	45.4	517	44.5	516	0.7
Female	54.4		55.5		
Percentage of patients(or caregivers) by education level					
No schooling	84		85.8		
Primary school (1 to 6 years of schooling)	6.8		7.5		
Intermediate school (7 to 9 years of schooling)	4.9	512	3.9	509	0.5
High school (10 to 12 years of schooling)	3		2.5		
Higher education (over 12 years of schooling)	0.8		0.2		
Percentage of patients(or caregivers) by socioeconomic status					
Very poor (having 0 to 5 household assets)	89		91		
Poor (having 6 to 9 household assets)	10.6	515	8.5	516	0.4
Not poor (having 10 to 14 household assets)	0.2		0.4		
Mean age of patients (in year)					
Male	15.9	230	16.5	240	0.7
Female	16.9	286	18.5	276	0.2

Table 5.4 Characteristics of health workers in RBF and non-RBF facilities, baseline – 2010

Variable	RBF facilities		Non-RBF facilities		
	%	N	%	N	P value
Percentage of health workers by sex					
Male	60	184	61	189	0.7
Female	40		39		
Percentage of health workers by positions assigned by the MOPH					
Doctor	18		17		
Nurse	19.8		16.6		
Midwife	24.5		24		
Assistant doctor	2	184	1	189	0.5
CHW Supervisor	4.7		7.8		
Vaccinator	30		30		
Pharmacist	0.5		2.6		
Percentage of health workers by work experience in the facilities surveyed					
Less than a year	8.8		11.4		
1 to 5 years	46.9	184	47.1	189	0.7
6 to 10 years	20.3		21.2		
Over 10 years	23.9		20.2		
Percentage of health workers who had interacted with a supervisor recently					
Within the past 30 days	84		83		
Within the past 31-90 days	12		9.3		
Within the past 3-6 months	2	184	2.6	189	0.2
More than 6 months	1		1		
Never	0.5		4.1		
Percentage of health workers who received training related to their jobs recently					
No training received	20		25		
1 - 7 days	33		30		
8 - 14 days	22	184	25	189	0.6
15 - 21 days	12		12		
Over 21 days	11		7		

The variables on training and supervision were created as categorical from the original continuous variables

Chapter 6: Overall effects

This chapter aims to examine RBF effects on patient satisfaction. The difference-in-differences (DID) estimation model described in chapter 4 will be employed. Using the DID estimator, the ordered logit model and linear probability model (LPM) will be used to detect RBF treatment effects on the outcome variables. The outcome variables are the 12 questions on patient satisfaction, which were asked during exit interviews with patients.

Data used in this section come from the exit interviews conducted with patients in health facilities. The questions were worded as: (i) How satisfied are you with travelling from your house to the health unit? (ii) How satisfied are you with the health facility cleanliness? (iii) How satisfied are you with how the health workers explained your illness? (iv) How satisfied are you with how the health workers explained your treatment? (v) How satisfied are you with the ease of getting the medicines the health workers prescribed? (vi) How satisfied are you with privacy during your visit? (vii) How satisfied are you with the amount of time the health worker spend with patient? (viii) How satisfied are you with the amount of time you spent waiting to be seen by a health provider? (ix) How satisfied are you with the respectfulness of the health providers? (x) How satisfied are with the cost of the visit to the health facility? (xi) How satisfied are you with the hours during which the health facility is open? and (xii) How satisfied are you with your overall visit? Responses were recorded on a four-point Likert scale: very dissatisfied=1, dissatisfied=2, satisfied=3, and very satisfied=4.

As described in chapter 4, clustering of data can happen at health facility level, as patients living in the catchment areas of health facilities share similar characteristics. To account for this, I adjusted standard errors at the level of health facilities to provide robust standard errors and statistical inferences.

6.1 Descriptive statistics

At first I present descriptive statistics on patient satisfaction responses in 2010. The purpose is to see variability of responses between the four-point scales, and decide about recoding of responses for further analysis. After recoding responses, aggregate indices on related items will be constructed and descriptive statistics will be provided to enable us to compare changes in the outcome between RBF and non-RBF groups over time.

Table 6.1 shows descriptive statistics on patient satisfaction responses in RBF and Non-RBF facilities in 2010. Table 6.1 shows that most variation in patients' responses occurred in the scales of "very satisfied" and "satisfied" – around 80% and more for most of the patient satisfaction items. For example, the percentage of patients who were overall "very satisfied" was 37%, and who were overall "satisfied" was 54% in RBF facilities. A similar pattern was seen in non-RBF facilities – with 41% for "very satisfied" and 53% for "satisfied" responses. Except for "satisfaction with traveling to health facilities", similar patterns are seen for other variables on patient satisfaction. Percentage of patient satisfaction with traveling to health facilities was 34% for "very satisfied", and 27% for "satisfied" responses in RBF facilities. In non-RBF facilities, it was 36% for "very satisfied", and 23% for "satisfied" responses.

Except for "satisfaction with traveling to health facilities", variation in the scales of "dissatisfied" and "very dissatisfied" contributed between 20% and 1%. For example, percentage of patients who were overall "dissatisfied" was 8%, and who were overall "very dissatisfied" was 1% in RBF facilities. A similar pattern with 6% for "dissatisfied" and 0% for "very dissatisfied" is seen in non-RBF facilities. Similar patterns of responses for "dissatisfied" and "very dissatisfied" responses are seen for nearly all items of patient satisfaction.

Table 6.1 also shows an interesting feature of patient satisfaction surveys, and that is a skewed distribution towards "satisfied" and "very satisfied" responses. This can indicate a positive bias. The frequency distribution in Table 6.1 resembles that one reported by Hansen et al. (2008) in Afghanistan. In the study by Hansen et al. (2008), in the frequency distribution on the eight items on overall patient satisfaction and perceived quality, it can clearly be seen that larger proportions of patient responses were on either satisfied or very satisfied responses which suggests a positive bias or overrating by the patients.

The tendency that patients overrate their levels of satisfaction has been well documented and explored in the literature (Peterson & Wilson, 1992; Williams, 1994). The skewness and positivity bias have been observed irrespective of using multi-items or single-item scales in satisfaction surveys, and are important practical issues to be considered during interpretation of survey findings (Peterson & Wilson, 1992). According to Peterson & Wilson (1992), the skewness and tendency could be (i) due to actual satisfaction with the service utilized or the performance experienced, or (ii) due to some methodological issues.

The methodological challenges are response rate bias, social desirability, use of positively worded question, order of questions, and interviews immediately after service provision. The response rate bias refers to the possibility that patients who are satisfied with services are more likely to participate; however, the patients who are dissatisfied may not participate in the study. In my study, the likelihood of such bias was minimized, as participants were selected prior to exit interviews and the response rate was higher than 95% in each of the three years.

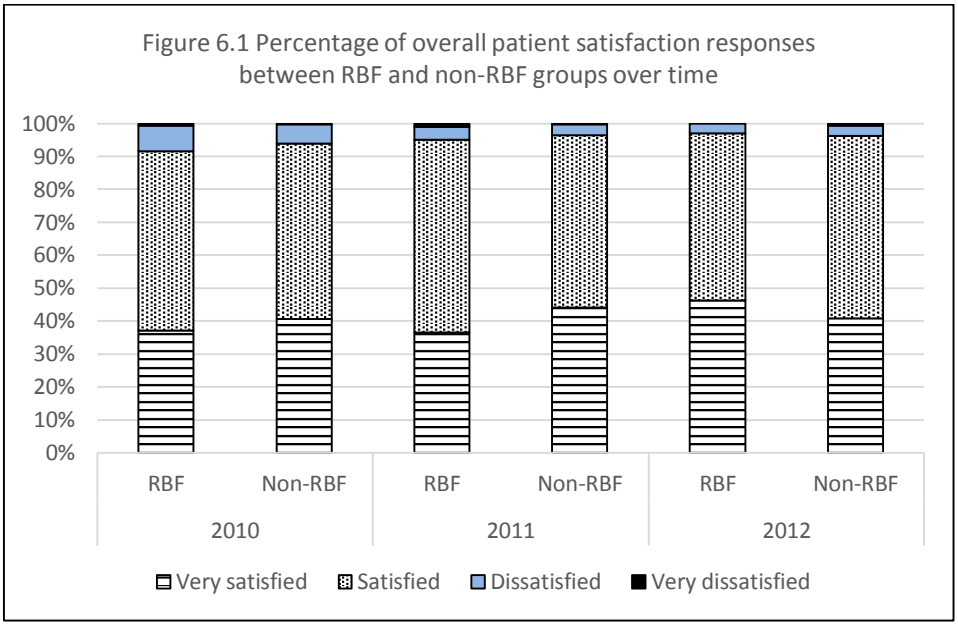
The social desirability bias occurs when patients express higher levels of satisfaction during face-to-face or telephone interviews in comparison with self-administered interviews. In my study it is highly likely that social desirability bias occurred, as face-to-face interviews were conducted with patients and as patients have more respect towards health providers in the Afghan context.

Use of positively worded questions can be perceived more positively by respondents, since this may influence patients' responses towards higher levels of satisfaction. In my study, all questions were positively worded (e.g., "How satisfied are you with ____?"). It was, therefore, likely that overrating on the levels of patient satisfaction occurred (i.e. more "very satisfied" and "satisfied" responses in comparison to "dissatisfied" and "very dissatisfied").

Order of questions can also influence respondents' levels of satisfaction. In a satisfaction survey, Peterson & Wilson (1992) found that asking an overall satisfaction question prior to other questions on specific aspect of quality slightly increases the respondents' tendency for a 'very satisfied' response. In my study, the "overall satisfaction" question was asked almost at the end of other questions, and this suggests that overrating of satisfaction may not have occurred due to ordering of questions.

Conducting interviews immediately after receiving services may have played significant role in influencing respondents to express higher levels of satisfaction in my study. However, due to having a randomized controlled trial design there seems to be little concern over the biasness of responses either in favour or against RBF facilities. Similar changes in patient responses can also be seen in 2011 and 2012. As an example, I provided variability of responses on the four-point scales on overall patient satisfaction (Figure 6.1).

Figure 6.1 shows that over 90% of responses occurred in the "very satisfied" and "satisfied" scales in 2010, 2011, and 2012. The same pattern is true for nearly all other items on patient satisfaction.

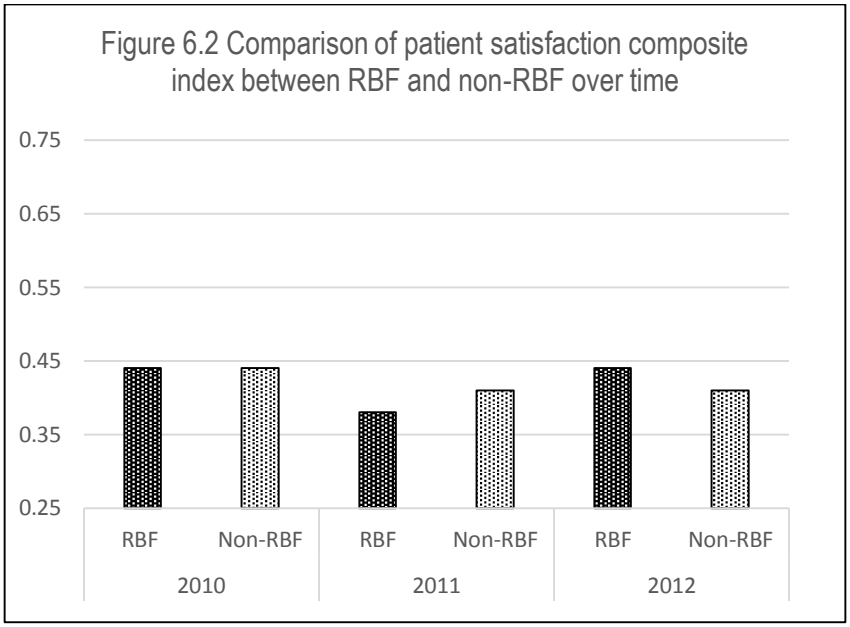


These changes show that variability between "very satisfied" and "satisfied" are much larger than those occurring between the "dissatisfied" and "very dissatisfied". This suggests that I can recode outcome variables as binary response: "very satisfied" equals one and "other three responses" equal zero.[40] After recoding the twelve items on patient satisfaction as binary responses, aggregate indices were constructed. The aggregate indices are weighted average of the items included in the respective index. For example, if a patient who was very satisfied with "sufficiency of visit time", and with "open hours of clinic", but not very satisfied with "waiting time",[41] the aggregate index which is the "responsive" index will be 0.67. This means on average the patient was 67% very satisfied with the three aspects of satisfaction. Similarly, a patient who was very satisfied with "respectful behaviour of providers"; but not very satisfied with "privacy in the clinic", the aggregate index which is the "assurance" index will be 0.5 which means on average the patient was 50% very satisfied with the two aspects of satisfaction. A composite index was also constructed, and this index is a weighted average of all the five indices.[42]

[40] In chapter 4, the marginal probability effects on the four scales also suggest recoding the outcome variables this way.
[41] Not very satisfied refers to either "satisfied", or "dissatisfied", or "very dissatisfied" because of recoding responses as binary
[42] The index is also a weighted average of all items, except the overall patient satisfaction.

Here the purpose of our descriptive statistics is to compare the changes in the outcome between RBF and non-RBF over time in comparison with the baseline differences. For the above purpose, Table 6.2 presents descriptive statistics. It shows that changes in the outcome for the five indices ranged from 1 to 4 percentage points between 2010 and 2011. The changes were not consistent – they were either in favour or against RBF facilities. For example, there was only 1 percentage point increase in the empathy index (from 28% in 2010 to 31% in 2011 in RBF, and from 31% in 2010 to 33% in 2011 in non-RBF); on the other hand, there was a 3 percentage points decline in the responsive index (from 35% in 2010 to 29% in 2011 in RBF, and from 35% in 2010 to 34% in 2011 in non-RBF). The largest and most inconsistent change relates to the assurance index where it decreased 4 percentage points in 2011 (from 53% in 2010 to 53% in 2011 in RBF, and from 54% in 2010 to 60% in 2011 in non-RBF), and it increased 4 percentage points in 2012 (from 53% in 2010 to 67% in 2012 in RBF, and from 54% in 2010 to 62% in 2012 in non-RBF).

Similar inconsistencies were also seen in other indices. For example, there was a decrease of 4 percentage points in expenses index in 2011 (from 68% in 2010 to 39% in 2011 in RBF, and from 66% in 2010 to 41% in 2011 in non-RBF), but there was an increase of 2 percentage points in 2012 in the expenses index (from 68% in 2010 to 42% in 2012 in RBF, and from 66% in 2010 to 42% in 2012 in non-RBF). The overall pattern of the changes is illustrated in Figure 6.2. Figure 6.2 shows that in comparison with the baseline, in 2011 patient satisfaction composite index decreased nearly 3 percentage points in RBF as compared with non-RBF. In 2012, however, patient satisfaction composite index increased nearly 3 percentage points in RBF than non-RBF in comparison with the index in 2010. In short, in comparison with the baseline, the changes in the outcome between RBF and non-RBF facilities over time seem to be inconsistent.



6.2 Treatment effects on patient satisfaction

We examined RBF treatment effects on different aspects of patient satisfaction over time by using the ordered logit model and LPM. I produced marginal probability effects at "very satisfied" scale of responses because most of the variations occurred between this scale and the other three scales of responses: satisfied, dissatisfied, and very dissatisfied (see chapter 4). In order to check consistency of results obtained by the ordered logit model, I also used the LPM.[43] The results from ordered logit model and LPM are provided in Table 6.3.

Table 6.3 shows that RBF treatment effects were not significant on any of the patient satisfaction aspects in 2011 and 2012. Results from the ordered logit model show that in 2011, the coefficients (or marginal effects) had negative sign on seven aspects of patient satisfaction. Results obtained by the LPM show that in 2011, the coefficients had negative sign on nine aspects of patient satisfaction. The seven aspects with negative marginal effects were: "overall satisfaction", "satisfaction with cleanliness of facilities", "satisfaction with respectfulness of providers", "satisfaction with privacy in facilities", "satisfaction with sufficiency of visit time", "satisfaction with adequacy of open hours", and "satisfaction with getting the prescribed medicine". In addition to them, the LPM shows negative coefficients on "satisfaction with cost of visiting facilities", and "satisfaction with waiting time".

[43] LPM will be used for the rest of multivariate analysis, as it allows the outcome to be either binary response or continuous variable (e.g., aggregate indices in our study).

In 2012, the only aspect with relatively higher impact, though not significant, relates to overall satisfaction. Results from the ordered logit model and LPM in 2012 show 9 percentage points and 8 percentage points increases in the proportion of patients who were very satisfied overall. Results in 2012 from the ordered logit model and LPM show that except for "satisfaction with cost of visiting facilities", and "satisfaction with getting the prescribed medicine" all other aspects of patient satisfaction were positively affected. Though the effects were not statistically significant.

6.3 Conclusion

From the findings in this chapter, one main conclusion is that this RBF program has weak effects on patient satisfaction on the whole. There was no impact on patient satisfaction in 2011. In 2012, there was some, but not significant, impact on patient satisfaction. The aspect with highest impact concerns overall satisfaction; however, it was not statistically significant. Moreover, the results obtained by the ordered logit model and LPM are similar in terms of magnitude and consistent in terms of signs of RBF treatment effects on patient satisfaction in 2011 and 2012.

Table 6.1 Percentage of responses on the four-point Likert scale in 2010

Variable	Group	Very satisfied	Satisfied	Dissatisfied	Very dissatisfied
Overall satisfaction	RBF	37	54	8	1
	Non-RBF	41	53	6	0
Satisfaction with travelling to health facilities	RBF	34	27	19	20
	Non-RBF	36	23	22	18
Satisfaction with cleanliness of health facilities	RBF	35	58	6	1
	Non-RBF	30	62	7	1
Satisfaction with respectfulness of providers	RBF	44	52	4	1
	Non-RBF	45	48	5	1
Satisfaction with the way providers explain illnesses	RBF	26	48	20	5
	Non-RBF	29	52	16	3
Satisfaction with the way providers explain treatment	RBF	29	52	16	3
	Non-RBF	33	51	14	1
Satisfaction with the cost of visiting health facilities	RBF	68	31	1	0
	Non-RBF	68	31	1	0
Satisfaction with privacy in health facilities	RBF	62	30	6	2
	Non-RBF	63	32	4	1
Satisfaction with sufficiency of visit time	RBF	36	54	9	1
	Non-RBF	36	51	12	0
Satisfaction with open hours of health facilities	RBF	37	53	9	1
	Non-RBF	35	56	9	1
Satisfaction with waiting time	RBF	32	40	22	7
	Non-RBF	34	44	16	6
Satisfaction with getting prescribed drugs	RBF	68	28	3	2
	Non-RBF	64	33	3	0

N= 517 for RBF, and N=506 for Non-RBF groups

Table 6.2 Aggregate indices showing average percentage of patients who were very satisfied[44]

Aggregate index	2010		2011		2012	
	RBF (n=517)	Non-RBF (n=506)	RBF (n=519)	Non-RBF (n=526)	RBF (n=496)	Non-RBF (n=472)
Composite	0.44	0.44	0.38	0.41	0.44	0.41
Empathy	0.28	0.31	0.31	0.33	0.30	0.31
Responsive	0.35	0.35	0.29	0.34	0.36	0.32
Assurance	0.53	0.54	0.53	0.60	0.67	0.62
Tangible	0.35	0.33	0.38	0.37	0.45	0.40
Expenses	0.68	0.66	0.39	0.41	0.42	0.42

Empathy contains "satisfaction with the way providers explain illness, and explain treatment", responsive contains "satisfaction with visit time, open hours, waiting time", assurance contains "satisfaction with respectfulness of providers, privacy", tangible contains "satisfaction with traveling to facilities, cleanliness of facilities", and expenses contains "cost of visit, getting prescribed drugs".
The aggregate indices are weighted average of the items included in the respective index. The items were recoded as binary: "very satisfied=1" and "other responses=0" before being included in the indices.

[44] RBF treatment effects on the aggregate indices are estimated in the next chapter.

Table 6.3 Marginal effects showing RBF treatment effects on various aspects of patients satisfaction

Variable	Ordered logit model			LPM			
	2011	2012	pseudo R^2	2011	2012	R^2	N
Overall satisfaction	-0.005	0.095	0.012	-0.021	0.081	0.019	2958
	(0.066)	(0.074)		(0.067)	(0.077)		
Satisfaction with travelling to health facilities	0.005	0.039	0.091	0.039	0.033	0.064	2949
	(0.054)	(0.056)		(0.064)	(0.067)		
Satisfaction with cleanliness of health facilities	-0.041	0.072	0.019	-0.023	0.055	0.035	2948
	(0.073)	(0.076)		(0.077)	(0.077)		
Satisfaction with respectfulness of providers	-0.047	0.045	0.026	-0.041	0.050	0.038	2958
	(0.071)	(0.079)		(0.075)	(0.077)		
Satisfaction with the way providers explain illnesses	0.051	0.065	0.009	0.035	0.033	0.004	2957
	(0.064)	(0.065)		(0.067)	(0.068)		
Satisfaction with the way providers explain treatment	0.050	0.049	0.006	0.048	0.024	0.007	2958
	(0.056)	(0.057)		(0.060)	(0.066)		
Satisfaction with cost of visiting health facilities	0.001	-0.014	0.051	-0.003	-0.015	0.078	2953
	(0.060)	(0.082)		(0.061)	(0.083)		
Satisfaction with privacy in health facilities	-0.018	0.078	0.004	-0.037	0.055	0.008	2957
	(0.073)	(0.092)		(0.073)	(0.089)		
Satisfaction with sufficiency of visit time	-0.018	0.0145	0.003	-0.010	0.012	0.007	2955
	(0.057)	(0.063)		(0.060)	(0.069)		
Satisfaction with open hours of health facilities	-0.046	0.089	0.009	-0.062	0.060	0.017	2956
	(0.063)	(0.088)		(0.068)	(0.085)		
Satisfaction with waiting time	0.007	0.089	0.010	-0.009	0.057	0.023	2955
	(0.051)	(0.059)		(0.059)	(0.064)		
Satisfaction with getting the prescribed medicine	-0.039	-0.019	0.051	-0.053	-0.028	0.086	2955
	(0.071)	(0.074)		(0.077)	(0.073)		

Standard errors are provided in parenthesis.

Chapter 7: Disaggregate effects

This chapter aims to explore RBF treatment effects on patient satisfaction between the three types of incentive administration mechanism. As discussed in chapter 3, RBF incentives were distributed among health workers according to four administration mechanisms: the salary-based, task-based, equal-amount, and mixed method.

In this chapter, I examine RBF treatment effect on patient satisfaction by three administration types. I dropped the health facilities under the mixed method administration because the sample size was small and because all the 8 facilities (4 RBF and 4 non-RBF) were located in Kandahar city. This does not allow for variability of information, even if I correct standard errors to account for the clustering of data within each health facility and between the 8 health facilities. Also the 6 health facilities (3 RBF and 3 non-RBF) included in the sample from Balkh province were dropped during analysis. The reason concerns the inconsistency in terms of administering incentives, as the RBF facilities in Balkh province changed from salary-based to task-based admin type in the last six months of the program. The change happened as AADA (NGO) took over all health facilities in Balkh province from CHA (another NGO) during that period of time (see Table 4.1).

In this chapter after presenting descriptive statistics, I used the LPM with DID estimation model, and estimated RBF treatment effects on: overall satisfaction, patient satisfaction composite score, and patient satisfaction indices such as empathy, responsive, assurance, tangible, and expenses for the whole sample, and for the sub-samples of salary-based, task-based, and equal-amount administrations. To address the problem of clustered data, standard errors were adjusted at the level of matched-pair facilities within each admin type, as well as for the whole sample. [45]

7.1 Descriptive statistics

For descriptive statistics, I present "patient satisfaction composite score" which is a weighted average of all items of patient satisfaction, except the overall satisfaction. [46] The results are provided in Table 7.1. Table 7.1 shows that in **2010** "patient satisfaction composite score"

[45] Please see Table 4.1 and related assumption on selecting non-RBF facilities as control groups for the matched-pair RBF facilities under the four administration types.
[46] Overall satisfaction was not included in the composite score as overall satisfaction was meant to measure a patient's satisfaction with all aspects of satisfaction as a whole.

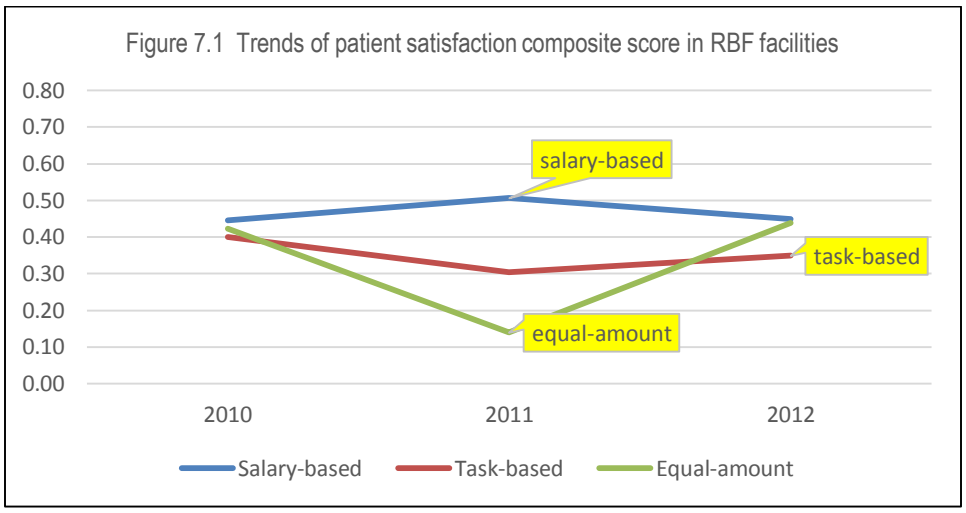
was slightly larger in RBF than non-RBF facilities under the salary-based and task-based mechanisms. But under the equal-amount, "patient satisfaction composite score" was smaller in RBF than non-RBF facilities: it was 0.43 in RBF and 0.48 in non-RBF.

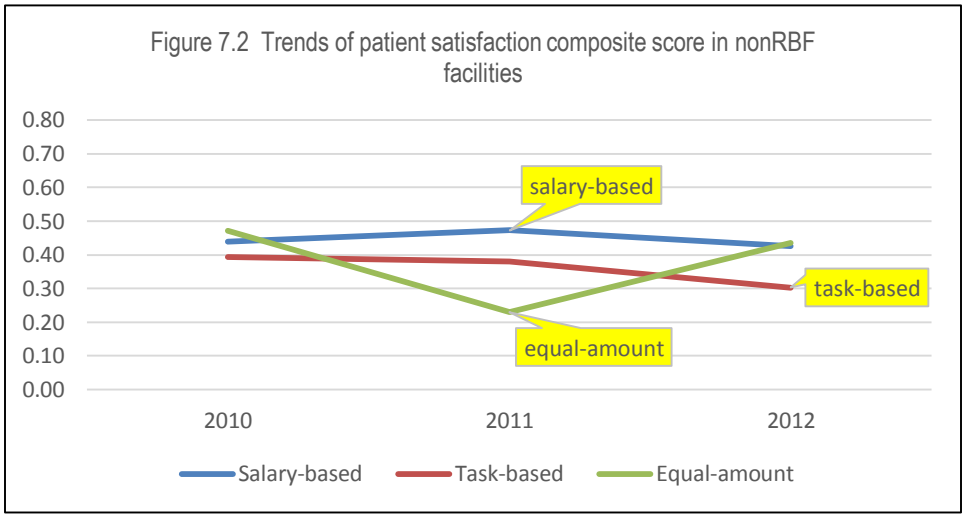
In **2011** the difference in "patient satisfaction composite score" was in favour of non-RBF in health facilities under the task-based and equal-amount administrations. For example, under the task-based administration, "composite score" was 0.31 in RBF and 0.38 in non-RBF facilities. This shows a decrease of 0.10 in RBF (from 0.41 in 2010 to 0.31 in 2011), and a decrease of 0.02 in non-RBF (from 0.40 in 2010 to 0.38 in 2011). So the lost was 0.08 or 8 percentage points for RBF in 2011. The largest difference was seen under the equal-amount admin– with the composite score of 0.14 in RBF and 0.23 in non-RBF facilities. This shows a decrease of 0.29 in RBF (from 0.43 in 2010 to 0.14 in 2011), and a decrease of 0.25 (from 0.48 in 2010 to 0.23 in 2011). So the loss was 0.04 or 4 percentage points. Unlike the above, under the salary-based administration there was a difference of 0.03 in favour of RBF (from 0.45 in 2010 to 0.52 in 2011 in RBF, and from 0.44 in 2010 to 0.48 in 2011 in non-RBF).

In **2012**, in comparison to 2010 and in comparison to non-RBF, the changes in "patient satisfaction composite score" were in favour of RBF. However, the changes were minor in facilities under the salary-based, and relatively higher in facilities under the other two administrations. For example, there was only 0.02 gain in the composite score under the salary-based admin (from 0.45 in 2010 to 0.46 in 2012 in RBF, and from 0.44 in 2010 to 0.43 in 2012 in non-RBF). There was a gain of 0.03 in health facilities under the task-based (from 0.41 in 2010 to 0.35 in 2012 in RBF, and from 0.40 in 2010 to 0.31 in 2012 in non-RBF). The gain in the composite score was 0.05 in health facilities under the equal-amount (from 0.43 in 2010 to 0.45 in 2012 in RBF, and from 0.48 in 2010 to 0.45 in 2012 in non-RBF).[47]

The above changes are also illustrated in Figures 7.1 and 7.2. Figures 7.1 and 7.2 also show that there was a *common pattern* in the changes between RBF and non-RBF facilities. This relates to the similarities of changes between RBF and non-RBF facilities under the specific administration type. This suggests that there may be considerable amounts of geographic or organization effect given that admin type is closely related to the implementing NGO.[48]

[47] The results might be different from those obtained by the DID estimation model as we have included control variables in the latter.
[48] At least that is the assumption for designing control groups for incentive admin types





7.2 RBF treatment effects on patient satisfaction

Both overall effects and disaggregated effects of RBF on overall satisfaction, composite score, and on patient satisfaction indices such as empathy, responsive, assurance, tangible, and expenses in 2011 and 2012 are provided in Table 7.2. In Table 7.2, RBF treatment effects will be compared between the findings from the whole sample, and the sub-samples.

RBF treatment effect on "overall satisfaction" shows that there was no significant effect in 2011 and 2012. Data from the whole sample, and data by the three admin types show that in 2011, the coefficients had negative signs. For example, it was – 0.021 for the whole sample, –

0.047 for the salary-based, – 0. 049 for the task-based, and – 0.040 for the equal-amount sub-samples. In 2012, except for the equal-amount, coefficients on overall satisfaction had positive sign. The largest coefficient was 0.10 for the sub-sample of task-based facilities, and 0.081 for the whole sample. However, none of them was statistically significant.

RBF treatment effect on "composite score" shows that there was no significant effect in 2011 and 2012. In 2011, the coefficients had negative sign when the whole sample data were used, and when data from the sub-samples for task-based, and for equal-amount were used. In 2012, the coefficients show some positive, though not significant, effect on the composite score.

RBF treatment effect on "empathy" shows that there was no significant effect in 2011 and 2012. In both years, the coefficients show that under the equal-amount administration, the effects were negative, though not significant. Under the salary-based, task-based, and for the whole sample, the coefficients show that in 2011 and 2012 RBF treatment effects were positive, though not significant. *RBF treatment effect on "responsive"* shows that there was no significant effect in either of the two years. In 2011, the coefficients show a negative effect under the sub-samples and for the whole sample. In 2012, treatment effects were positive, but not significant.

RBF treatment effect on "assurance" shows that there was a significant effect health facilities under the task-based administration in 2011. In 2011, there was a reduction of 21 percentage points as an average for two items (satisfaction with respectfulness of providers, and satisfaction with privacy) in health facilities under the task-based admin ($p=0.031$). There was a negative effect when treatment effect was estimated for the whole sample; but the effect was not significant. In 2012, the coefficients show positive, but not significant, effect.

RBF treatment effect on "tangible" shows that all coefficients in 2011 and 2012 indicate positive, though not significant, effect. Unlike this, *RBF treatment effect on "expenses index"* shows negative effect in 2011 and 2012 under the task-based and for the whole sample. In 2011, the results from salary-based and equal-amount admin types show positive effect, though not significant. In 2012, except for the equal-amount admin type, the coefficients obtained from the whole sample, and under the salary-based and task-based administrations show negative effects. However, none of the effects were significant.

7.3 Conclusions

Consistent with my findings in previous chapter, here too it was found that RBF did not have a positive effect on patient satisfaction over a one year period (in 2011). Over a two-year period (in 2012) there was some positive, but not significant, effect on patient satisfaction in health facilities. This was true for the sub-samples of salary-based, task-based, and equal-amount administrations. This suggests that paying financial incentives may not have the power to positively affect patient satisfaction; no matter how the incentives are distributed among health workers at the health facilities.

A possible explanation can relate to the design of RBF where except for the financial incentives, other support such as non-financial incentives and administrative and logistical support were kept constant between RBF and non-RBF facilities over time. The role of both financial and non-financial incentives has been reported as key factors in the success of RBF and similar approaches (Van Herck et al., 2010; Oxman & Fretheim, 2009).

Table 7.1 composite index satisfaction score in RBF and non-RBF facilities

Variable	Admin type	2010		2011		2012	
		RBF	Non-RBF	RBF	Non-RBF	RBF	Non-RBF
Composite score	Salary-based	0.45	0.44	0.52	0.48	0.46	0.43
	Task-based	0.41	0.40	0.31	0.38	0.35	0.31
	Equal-amount	0.43	0.48	0.14	0.23	0.45	0.45

Composite score is a weighted average of all items on patient satisfaction, except the overall satisfaction. The responses for each items were recoded as binary response.

Table 7.2 RBF treatment effects on aggregate indices of patient satisfaction under the 3 admin types

Outcome variable	All RBF		Salary-based		Task-based		Equal-amount	
	2011	2012	2011	2012	2011	2012	2011	2012
Overall satisfaction	-0.021	0.081	-0.047	0.046	-0.049	0.102	-0.040	-0.030
	(0.067)	(0.077)	(0.097)	(0.121)	(0.127)	(0.108)	(0.185)	(0.231)
Composite score	-0.009	0.029	0.036	0.022	-0.062	0.037	-0.023	0.047
	(0.045)	(0.049)	(0.062)	(0.078)	(0.073)	(0.081)	(0.115)	(0.125)
Empathy	0.042	0.029	0.083	0.025	0.129	0.072	-0.237	-0.193
	(0.061)	(0.065)	(0.088)	(0.119)	(0.105)	(0.106)	(0.133)	(0.148)
Responsive	-0.028	0.043	-0.004	0.014	-0.082	0.078	-0.070	0.027
	(0.049)	(0.060)	(0.076)	(0.105)	(0.087)	(0.087)	(0.135)	(0.168)
Assurance	-0.038	0.053	0.083	0.021	-0.210*	0.100	0.077	0.133
	(0.063)	(0.073)	(0.085)	(0.114)	(0.093)	(0.120)	(0.157)	(0.214)
Tangible	0.008	0.045	0.016	0.071	0.014	0.023	0.021	0.143
	(0.061)	(0.061)	(0.089)	(0.079)	(0.086)	(0.103)	(0.146)	(0.144)
Expenses	-0.029	-0.021	0.004	-0.022	-0.163	-0.090	0.092	0.124
	(0.059)	(0.068)	(0.076)	(0.109)	(0.110)	(0.089)	(0.162)	(0.143)

p*<0.05

Chapter 8: Effects on intermediary factors

In this chapter we are interested in examining RBF treatment effects on the intermediate factors which may have a positive impact on patient satisfaction. More specifically I investigate as whether RBF has any effect on health provider performance; on availability of medicine, vaccines, equipment; and on physical appearance and cleanliness of health facilities. These factors are instrumental in the process of healthcare delivery and quality improvement that can influence patient satisfaction.

The chapter is organized into three sections. *Section one* covers RBF treatment effects on health provider performance. I measure health provider performance by studying the patient – provider interaction. Patient – provider interaction refers to the implicit and explicit procedural standards that a health worker is supposed to perform in order to build rapport with a patient and provide health services to the patient. To measure patient – provider interaction I utilize data collected during patient visit. In each health facility, the interaction between patients and health providers was observed for up to 10 patients.[49] The purpose was to identify whether the health providers were following the procedural standards during their interactions with patients. These standards mainly entail communication skills and some basic, but vital, clinical procedures.

There were 13 questions to measure patient – provider interaction. These questions were: (i) whether the health provider greets the patient, (ii) asks patient age, (iii) asks about chief complaints of the patient (or the reasons for visiting the provider), (iv) asks about the duration of the illness, (v) asks about the past treatment, (vi) examines the patient, (vii) tells the name of the illness to the patient, (viii) discusses the course of the illness to the patient, (ix) advises the patient about home care, (x) advises the patient on how to take the medicine, (xi) tells the patient about the side effects of the prescribed drug(s), (xii) tells the patient about follow-up visits, (xiii) and asks the patient whether he/she has any questions.

[49] Exit interviews were conducted with the same patients by a second surveyor from the survey team.

In our analysis, the 13 questions, which had binary responses, were categorized into three indices: history taking, medical examination, and medical advice. A composite score which is a weighted average of the three indices was also computed. History taking is composed of: the health provider greets the patient, asks patient age, asks about chief complaints, asks about the duration of the illness, and asks about the past treatment. Medical examination includes items such as the health provider examines the patient, tells the name of the illness, and discusses the course of the illness. Medical advice contains items such as the health provider explains about home care, discusses how to take the prescribed medicine, discusses the side effects of the medicine, talks about follow-up visits, and asks the patient if she/he has any questions. Each index is computed as the weighted average score for the related items. The control variables included in the model are the same used in the last two chapters.

Section two reports RBF treatment effects on the availability of essential drugs, vaccines, and equipment. Data used here come from direct observation of pharmacy, vaccination room, laboratory, and stock records and registration in health facilities. Availability and functionality of these important resources was recorded as "yes", otherwise "no". For example a microscope which was not working properly was recorded as "no" in the survey questionnaire, or availability of an expired (or near to expire) drug was recorded as "no" or not available at the time of observation. In my study I computed one composite score, and three indices which are weighted average scores on availability of several types of essential drugs, vaccines, and equipment. The composite score is a weighted average score of the three indices. The composite score, and indices were used as outcome variables in DID estimation model. The control variables are: distance of a health facility from the provincial capital, and type of health facility (whether the facility was a comprehensive health centre or otherwise).

Section three deals with RBF treatment effects on physical appearance and cleanliness of health facilities (e.g. good physical condition, and cleanliness and tidiness of health facilities). Data used here come from direct observation of health facilities. To collect data for physical condition, a surveyor walked around the facility and observed the condition of the windows, doors, interior walls, exterior walls, roof, and the ground and fences. The surveyor recorded the score "1" if the subject in question did not need repair, and recorded the score "0" otherwise. Similarly, for the cleanliness of a health facility, a surveyor walked around the facility and observed several rooms (e.g. delivery room), sections (e.g. pharmacy), toilets, and grounds. The surveyor recorded the score "1" if the area in question was clean, and

recorded the score "0" otherwise. For my analysis here, I computed weighted average score for "good physical condition", for "cleanliness and tidiness" of health facilities, and for "compose score" and used them as outcome variables. The control variables are the same used in section two of the chapter. Consistent with my approach in previous chapters, first descriptive statistics are provided, and then the LPM with the DID estimation model will be used for each section.[50]

8.1 Health provider performance

For the ***descriptive statistics*** I used the composite score which is an overall average of all the 13 items related to health provider performance, measured through patient – provider interaction. The results are provided in Table 8.1. Table 8.1 shows that in **2010** there were minor to no differences between RBF and non-RBF facilities. For example, composite score was 0.50 in RBF and 0.51 in non-RBF facilities. Using the sub-samples, I see that composite score was 0.61 in RBF and 0.59 in non-RBF facilities under the equal-amount administration. Under the other two types, there was no difference in the composite score of health provider performance between RBF and non-RBF.

In **2011**, unlike the equal-amount administration where composite score decreased, composite score increased in RBF and non-RBF facilities as compared to 2010. However, the gains were not impressive. Using data from the whole sample, I see that there was a gain of 0.02 or 2 percentage points (from 0.50 in 2010 to 0.54 in 2011 in RBF, and from 0.51 in 2010 to 0.53 in 2011 in non-RBF). A similar picture was seen in health facilities under the salary-based and task-based administrations. For example, there was a gain of 0.02 in salary-based (from 0.50 in 2010 to 0.56 in 2011 in RBF, and from 0.50 in 2010 to 0.54 in 2011 in non-RBF). Under the task-based the gain was only 1 percentage point (from 0.49 in 2010 to 0.54 in 2011 in RBF, and from 0.49 in 2010 to 0.53 in 2011 in non-RBF). In health facilities under the equal-amount administration, on the other hand, there was a loss of 5 percentage points (from 0.61 in 2010 to 0.43 in 2011 in RBF, and from 0.59 in 2010 to 0.46 in 2011 in non-RBF).

In **2012**, a somewhat similar picture can be painted. For example, there was only 2 percentage points gain as a whole in the composite score (from 0.50 in 2010 to 0.52 in 2012 in RBF, and from 0.51 in 2010 to 0.51 in 2012 in non-RBF). In health facilities under the salary-based

[50] In case the outcome is continuous variable, the same LPM is used, and is called OLS (ordinary least square).

administration there was only 1 percentage point gain (from 0.50 in 2010 to 0.54 in 2012 in RBF, and from 0.50 in 2010 to 0.53 in 2012 in non-RBF). Similarly in health facilities under the task-based there was a gain of 2 percentage points (from 0.49 in 2010 to 0.50 in 2012 in RBF, and from 0.49 in 2010 to 0.48 in 2012 in non-RBF). Unlike the above, in health facilities under the equal-amount there was a loss of 2 percentage points in the composite score (from 0.61 in 2010 to 0.55 in 2012 in RBF, and from 0.59 in 2010 to 0.55 in 2012 in non-RBF). The results from the descriptive statistics suggest no considerable differences in health provider performance between RBF and non-RBF over a one year and two years period.

For the ***RBF treatment effects***, I use the DID estimation model and present treatment effects (Table 8.2). Table 8.2 shows that RBF did not have significant effect on composite score, history taking, medical examination, and medical advice over time.

RBF treatment effect on "*composite score"* was positive as a whole in 2011.[51] However, when data from the sub-samples were used, negative coefficients were seen in health facilities under the task-based and equal-amount administrations. In 2012, all the coefficients had positive sign, indicating positive treatment effect on composite score. RBF treatment effects, however, were not significant either in 2011 or 2012.

RBF treatment effect on *"history taking*" was negative only in health facilities under the task-based administration in 2011. As a whole, RBF treatment effect was positive in 2011. In 2012, all the coefficients show positive treatment effect on "history taking" of health providers. Neither in 2011 nor in 2012 was the treatment effect significant, though.

RBF treatment effect on "*medical examination"* was negative in health facilities under the salary-based and equal-amount administrations. As a whole I found a positive, though not significant, treatment effect on "medical examination" conducted by health providers in 2011. A similar pattern was seen in 2012: a negative, but not significant, effect of RBF on "medical examination" in facilities under the salary-based and equal-amount administrations. As a whole, the results show a positive, but not significant, treatment effect on "medical examination" in 2012.

[51] Whenever "as a whole is used" it refers to the results obtained from the data of the whole sample.

RBF treatment effect on "*medical advice*" was negative in health facilities under the task-based and equal-amount administrations in 2011. As a whole, there was a positive, but not significant, treatment effect in 2011. In 2012, a similar pattern was seen. Treatment effects were negative under the task-based and equal-amount administrations. Under the salary-based, and as a whole treatment effect was positive, though not significant in 2012.

The results obtained from the DID estimation model suggest that in general RBF positively affected health provider performance in 2011 and 2012. However, the effect was not significant. Data from the sub-samples provide results that show an inconsistent pattern on RBF treatment effect on health provider performance. This suggests that there was no clear difference between the three admin types in 2011 and 2012.

8.2 Availability of medicines, vaccines, and equipment

For the ***descriptive statistics*** I computed composite score which is an average score of all items related to availability of medicine, vaccines, and equipment in RBF and non-RBF facilities (Table 8.3). Table 8.3 shows that there was some differences in composite score between RBF and non-RBF facilities in **2010**. Data from the whole sample show that composite score was slightly higher in RBF than non-RBF facilities (0.76 versus 0.74). Data from the sub-samples show inconsistency across the admin types. For example, under the salary-based, composite score was 0.75 in RBF and 0.75 in non-RBF facilities; but under the other two administrations composite score was higher in RBF facilities. The composite score under the task-based administration was 0.73 in RBF and 0.71 in non-RBF, and under the equal-amount the score was 0.80 in RBF and 0.72 in non-RBF facilities. In **2011**, data from the whole sample show that there was a slight loss (1 percentage point) in the composite score (from 0.76 in 2010 to 0.78 in 2011 in RBF, and from 0.74 in 2010 to 0.77 in 2011 in non-RBF). Data from the sub-samples show inconsistency across the admin types. For example, under the salary-based there was 1 percentage point increase in the composite score (from 0.75 in 2010 to 0.79 in 2011 in RBF, and from 0.75 in 2010 to 0.78 in 2011 in non-RBF); under the task-based there was no gain or loss (from 0.73 in 2010 to 0.79 in 2011 in RBF, and from 0.71 in 2010 to 0.77 in 2011 in non-RBF). Unlike the above, under the equal-amount sub-sample there was an increase of 16 percentage points (from 0.80 in 2010 to 0.66 in RBF, and from 0.72 in 2010 to 0.64 in 2011 in non-RBF).

In **2012**, loss in the composite score was seen, at varying extents, when data from the whole sample and sub-samples were used. For example, the whole sample data show that there was a loss of 4 percentage points (from 0.76 in 2010 to 0.76 in 2012 in RBF, and from 0.74 in 2010 to 0.78 in 2012 in non-RBF). Under the salary-based admin, the loss was 2 percentage points (from 0.75 in 2010 to 0.78 in 2012 in RBF, and from 0.75 in 2010 to 0.80 in 2012 in non-RBF), and under the task-based the loss was 4 percentage points (from 0.73 in 2010 to 0.71 in 2012 in RBF, and from 0.71 in 2010 to 0.73 in 2012 in non-RBF). The largest loss was seen in health facilities under the equal-amount with 17 percentage points decrease in the composite score (from 0.80 in 2010 to 0.79 in 2012 in RBF, and from 0.72 in 2010 to 0.88 in 2012 in non-RBF). The largest loss under the equal-amount is contrary to the highest gain in the composite score under the same admin type in 2011.

The descriptive statistics suggest that in general there were some losses in the composite score in health facilities over time. Looking into the sub-samples, I also see some inconsistencies in the composite score between the three admin types. The admin type with the most inconsistent pattern relates to the equal-amount administration.

For the ***RBF treatment effects*** I provide treatment effects on availability of medicine, vaccines, and equipment from the DID estimation model (Table 8.4). Table 8.4 shows that except for the composite score, and availability of equipment in health facilities under the equal-amount administration, there was no significant treatment effect on composite score, on availability of medicine, vaccines, and equipment over time. Under the equal-amount administration, the treatment effect on composite score was negative and significant in 2012. Also in 2011 and 2012, treatment effect was negative and significant on availability of equipment in health facilities under the equal-amount administration.

Considering composite score as an average score for all items related to availability of equipment, medicine, and vaccines, it was found that in 2011 "*composite score"* was negatively affected as a whole. The only exception was health facilities under the salary-based administration. In 2012, treatment effect was negative for all health facilities, which included the health facilities in the three sub-samples. Except for the equal-amount, treatment effect was not significant either in 2011 and 2012. In the equal-amount facilities, only in 2012 the effect was significant with a decrease of 17 percentage points in the composite score ($p=0.025$). As a whole, RBF treatment effect was negative in 2011 and 2012, but the effect

was not significant. When it comes to the sub-samples data, there some consistencies in terms of treatment effect across the admin types. The admin type with significantly negative impact relates to the equal-amount facilities.

RBF treatment effect on "*availability of medicine*" was positive as a whole in 2011. The same was true for the sub-samples of salary-based and task-based administrations in 2011. For the equal-amount sub-sample, RBF treatment effect was negative in 2011. In 2012, except for health facilities under the salary-based, coefficients of treatment effect had negative sign, indicating negative impact. However, none of the effects were significant in 2011 and 2012.

RBF treatment effect on "*availability of vaccines*" was negative as a whole in 2011. Data from sub-samples show that RBF treatment effect was positive only in health facilities under the equal-amount administration in 2011. In 2012, RBF treatment effect was positive only in health facilities under the task-based administration. Data from the whole sample show a negative effect in 2012. These effects, however, were not significant in either years.

RBF treatment effect on "availability of equipment" was negative and significant in health facilities under the equal-amount administration in 2011 and 2012. RBF treatment effect was negative as a whole, and in health facilities under the task-based and equal-amount administration. Under the equal-amount administration, there was a decrease of 20 percentage points in the availability of equipment in 2011 ($p=0.010$). In 2012, RBF treatment effect was negative as a whole, and when sub-samples data were analysed. However, the effect was only significant under the equal-amount administration where there was a decrease of 23 percentage points in the availability of equipment ($p=0.031$).

The above results suggest that as a whole RBF had some negative, but not significant, effects on availability of medicine, vaccines, and equipment in 2011 and 2012. Looking to the data from the three admin types, it was found that availability of equipment was negatively and significantly affected in health facilities under the equal-amount administration. However, there were considerable inconsistencies in the treatment effects on availability of medicine and vaccines across the three admin types. This suggests that except for the availability of equipment, there were no clear differences between the three admin types in terms of RBF treatment impact on availability of essential drugs and vaccines.

8.3 Physical condition and cleanliness of health facilities

For the ***descriptive statistics*** we provided composite score which is an average score of all items related to physical condition and cleanliness of health facilities (Table 8.5). Table 8.5 shows that as a whole there was no difference between RBF and non-RBF in terms of physical condition and cleanliness of health facilities in **2010**. Data from the sub-samples, however, show considerable differences between RBF and non-RBF facilities. The differences were not consistent across the three admin types. For example, data from the whole sample show that composite score was 0.62 in RBF and 0.61 in non-RBF, and data from the sub-samples show that under the salary-based the composite score was 0.66 in RBF and 0.76 in non-RBF, and under the task-based it was 0.63 in RBF and 0.55 in non-RBF, and under the equal-amount administration it was 0.49 in RBF and 0.43 in non-RBF facilities.

In **2011**, as a whole there was some gains in the composite score. However, some inconsistencies in the gains or losses in the composite score can be seen in the sub-sample of health facilities. For example, data from the whole sample show a gain of 4 percentage points in 2011 (from 0.62 in 2010 to 0.72 in 2011 in RBF, and from 0.61 in 2010 to 0.67 in 2011 in non-RBF); and data from the sub-samples show a loss of 2 percentage points in the salary-based facilities (from 0.66 in 2010 to 0.68 in 2011 in RBF, and from 0.76 in 2010 to 0.80 in 2011 in non-RBF); a gain of 9 percentage points (from 0.63 in 2010 to 0.75 in 2011 in RBF, and from 0.55 in 2010 to 0.58 in 2011 in non-RBF) in the task-based; and a gain of 4 percentage points (from 0.49 in 2010 to 0.61 in 2011 in RBF, and from 0.43 in 2010 to 0.51 in 2011 in non-RBF) in the equal-amount facilities.

In **2012**, as a whole there was some gains. But when it comes to the sub-samples, except for the health facilities under the task-based, there were some to considerable increases in the composite score in the salary-based and equal-amount administrations. For example, as a whole there was a gain of 2 percentage points (from 0.62 in 2010 to 0.71 in 2012 in RBF, and from 0.61 in 2010 to 0.68 in 2012 in non-RBF). However, under the task-based there was a decline of 2 percentage points (from 0.63 in 2010 to 0.70 in 2012 in RBF, and from 0.55 in 2010 to 0.59 in 2012 in non-RBF). The pattern was different under the other two admin types: under the salary-based there was an increase of 5 percentage points (from 0.66 in 2010 to 0.77 in 2012 in RBF, and from 0.76 in 2010 to 0.82 in 2012 in non-RBF), and under the equal-amount the gain was 22 percentage points (from 0.49 in 2010 to 0.71 in 2012 in RBF,

and from 0.43 in 2010 to 0.42 in 2012 in non-RBF). The results from the descriptive statistics suggest that in general, there were some gains. But when it comes to the results from the sub-samples there were considerable differences and inconsistencies between the three admin types.

For the ***RBF treatment effects*** I present the results from the DID estimation model (Table 8.6). Table 8.6 shows that there was no significant treatment effect of RBF on composite score of facilities physical condition and cleanliness, on physical condition and cleanliness of health facilities separately.

RBF treatment effect on "*composite score*" of facilities physical condition and cleanliness was negative only in health facilities under the salary-based administration in 2011. As a whole, as well as in the sub-sample of task-based and equal-amount administrations, RBF treatment effect was positive, though not significant, in 2011. In 2012, except for the task-based administration, in all other facilities RBF treatment effect was positive. In general, RBF treatment effect was positive in 2011 and 2012. Looking into the sub-samples, however, inconsistencies were seen with regard to the treatment effect on the composite score of facilities physical condition and cleanliness. None of the effects were significant, however.

RBF treatment effect on "*good physical condition*" of health facilities was not significant in either of the two years. In 2011, except for health facilities under the salary-based, all coefficients show positive effect of the intervention. In 2012, all the coefficients indicate positive treatment effect. However, neither in 2011 nor in 2012 were the effects significant.

RBF treatment effect on "*cleanliness*" of health facilities was not significant in 2011 and 2012. In 2011, as a whole the effect was negative. It was also negative in the salary-based administration. Under the other two admin types, treatment effect was positive in 2011. In 2012, except for the equal-amount administration, all coefficients, including the one from the whole sample, show negative treatment effect. In general, RBF treatment effect was negative in 2011 and 2012; however, the effect was not significant in either of the years.

8.4 Conclusions

In section one, we found that RBF had some positive, but not significant, effects on health provider performance. My exploration between the three admin types found that RBF effects on health provider performance did not differ much between the three administration types.

In section two, I found that in general RBF had some negative, though not significant, effects on availability of medicine, vaccines, and equipment. When it came to my exploration between the three admin types, it was found that availability of equipment was negatively and significantly effected in the equal-amount facilities in 2011 and 2012. Even if this shows that the equal-amount facilities did not have sufficient equipment, I cannot rule out the possibility of sample size effect. This means I had data from a small sample under the equal-amount administration, and since the unit of analysis for availability of medicine, vaccines, and equipment is health facilities, the sample remains 46 for a panel of RBF and non-RBF facilities over the three data points. Therefore, it is possible that RBF treatment effect on availability of equipment is overestimated under the equal-amount facilities.

In section three, I found that in general RBF had some positive, but not significant, effects on physical condition of health facilities. However, when it came to RBF treatment effects on cleanliness of health facilities, it was found that in general RBF had some negative, though not significant, impact on this aspect of quality. My exploration on the treatment effects between the three admin types found some inconsistencies between the three administrations. This and the conclusions from sections one and two suggest that no difference can be expected in the RBF treatment effects between the three administration types.

In this chapter, it was found that RBF affected health provider performance, availability of essential drugs, vaccines, equipment, and physical appearance of health facilities. However, the effects were not significant. In summary, the evidence drawn in this chapter suggests that monetary incentives alone – no matter how they are administered at health facilities – in a RBF program may not be sufficient to affect determinants of patient satisfaction at health facility level.[52]

[52] Determinants of patient satisfaction at health facility level refer to health provider performance, availability of medicine, vaccines, equipment, as well as physical appearance of health facilities.

Table 8.1 Composite score on health provider performance in RBF and non-RBF facilities

Sample	2010		2011		2012	
	RBF	non-RBF	RBF	non-RBF	RBF	non-RBF
All facilities	0.50	0.51	0.54	0.53	0.52	0.51
Salary-based	0.50	0.50	0.56	0.54	0.54	0.53
Task-based	0.49	0.49	0.54	0.53	0.50	0.48
Equal-amount	0.61	0.59	0.43	0.46	0.55	0.55

Composite score is a weighted average of all the 14 items of patient – provider interaction

Table 8.2 RBF treatment effects on health provider performance

Outcome variable	The whole sample n = 2958		Salary-based n = 1349		Task-based n = 960		Equal-amount n = 413	
	2011	2012	2011	2012	2011	2012	2011	2012
Composite score	0.027	0.020	0.036	0.022	-0.062	0.036	-0.023	0.047
	(0.035)	(0.031)	(0.062)	(0.078)	(0.072)	(0.081)	(0.115)	(0.125)
History taking	0.004	0.031	0.005	0.002	-0.018	0.043	0.019	0.053
	(0.032)	(0.027)	(0.054)	(0.044)	(0.053)	(0.044)	(0.075)	(0.085)
Medical exam	0.048	0.013	-0.028	-0.050	0.093	0.053	-0.079	-0.049
	(0.053)	(0.055)	(0.092)	(0.102)	(0.073)	(0.067)	(0.134)	(0.145)
Medical advice	0.031	0.017	0.066	0.063	-0.014	-0.033	-0.041	-0.060
	(0.043)	(0.040)	(0.070)	(0.054)	(0.049)	(0.060)	(0.085)	(0.111)

History taking index includes: providers greeting patients, asking their age, asking about the reason for the visit, asking about past illnesses and treatments. Medical examination index includes: providers examining patients, telling them name of their illness, and explaining course of their illness. Medical advice index includes: providers explaining about home care, discussing how to take the medicine, discussing about side effects, talking about follow-up visits, and asking patients for any questions. Each index was computed as mean of the items included in the index. Control variables included are the same used for patient satisfaction.
Standard errors, which are provided in parenthesis, were adjusted at the level of health facilities.

Table 8.3 Composite score on availability of medicine, vaccines, and equipment

Sample	2010		2011		2012	
	RBF	non-RBF	RBF	non-RBF	RBF	non-RBF
The whole sample	0.757	0.736	0.783	0.767	0.764	0.779
Salary-based	0.747	0.751	0.789	0.781	0.784	0.797
Task-based	0.730	0.714	0.789	0.772	0.710	0.726
Equal-amount	0.799	0.724	0.663	0.635	0.789	0.882

Composite score is a weighted average of all questions related to availability of essential drugs, vaccines, and equipment in health facilities. Availability of these resources was recorded as "yes: available" if they were in functional order (e.g., any drugs with expired date was recorded as "not available").

Table 8.4 RBF treatment effects on availability of medicine, vaccines, and equipment

Variable	The whole sample n = 324		Salary-based n = 129		Task-based n = 108		Equal-amount n = 46	
	2011	2012	2011	2012	2011	2012	2011	2012
Composite score	-0.008	-0.043	0.007	-0.006	-0.008	-0.042	-0.046	-0.171*
	(0.030)	(0.033)	(0.038)	(0.029)	(0.055)	(0.055)	(0.054)	(0.069)
Availability of medicine	0.008	-0.048	0.015	0.031	0.032	-0.108	-0.010	-0.135
	(0.042)	(0.049)	(0.068)	(0.059)	(0.074)	(0.095)	(0.072)	(0.110)
Availability of vaccines	-0.021	-0.040	-0.036	-0.037	-0.001	0.018	0.075	-0.152
	(0.042)	(0.044)	(0.022)	(0.046)	(0.092)	(0.083)	(0.146)	(0.113)
Availability of equipment	-0.015	-0.059	0.042	-0.014	-0.065	-0.040	-0.204**	-0.226*
	(0.044)	(0.041)	(0.065)	(0.053)	(0.082)	(0.071)	(0.069)	(0.095)

p*<0.05, p**<0.01
Standard errors, which are provided in parenthesis, were adjusted at the health facility level.

Table 8.5 Composite score on physical condition and cleanliness of health facilities

Sample	2010		2011		2012	
	RBF	non-RBF	RBF	non-RBF	RBF	non-RBF
The whole sample	0.616	0.607	0.721	0.670	0.709	0.675
Salary-based	0.655	0.762	0.683	0.801	0.770	0.818
Task-based	0.634	0.548	0.748	0.578	0.654	0.591
Equal-amount	0.492	0.432	0.606	0.511	0.705	0.416

Composite score is a weighted average of "physical condition" and "cleanliness" of health facilities. "Good physical condition" of a health facility is an average score of six questions related to functionality and appearance of the facility, and "cleanliness" of a health facility is an average score of ten questions related to cleanliness of different rooms/parts of the facility. The above information was collected by direct observation.

Table 8.6 RBF treatment effects on physical condition and cleanliness of health facilities

Variable	The whole sample n = 324		Salary-based n = 129		Task-based n = 108		Equal-amount n = 46	
	2011	2012	2011	2012	2011	2012	2011	2012
Composite score	0.034	0.006	-0.016	0.054	0.073	-0.038	0.036	0.206
	(0.064)	(0.072)	(0.096)	(0.093)	(0.122)	(0.135)	(0.114)	(0.179)
Good physical condition of health facilities	0.056	0.009	-0.012	0.068	0.102	0.002	0.025	0.047
	(0.080)	(0.081)	(0.120)	(0.106)	(0.149)	(0.126)	(0.196)	(0.220)
Cleanliness of health facilities	-0.022	-0.017	-0.096	-0.006	0.044	-0.078	0.046	0.353
	(0.073)	(0.089)	(0.098)	(0.118)	(0.144)	(0.192)	(0.141)	(0.201)

Standard errors, which are shown in parenthesis, were adjusted at health facilities level.
The OLS model with DID estimation was used.

Chapter 9: Discussion and policy implications

This chapter discusses the major findings of this study. The strengths and limitations of the study will be highlighted, and the implications of the study in relation to influencing policy decisions in the health sector in a post conflict and fragile setting such as Afghanistan will be discussed. Areas for future investigations will be explored and recommended.

9.1 Results discussion

Throughout this study, we have found that in general RBF did not have any impact on patient satisfaction with healthcare delivery in the first year of implementation. In the second year, however, there was some positive but not significant impact of RBF on overall patient satisfaction. With respect to the above change, assuming the intended objective is to improve patient satisfaction, two critical arguments emerge. The arguments aim to help policy makers decide on the design and implementation of a more patient-centred approach for next generation of performance-based financing depending on the availability of donor funding and political will of the MOPH. The first argument concerns the positive effect of RBF on overall satisfaction in relation to the amounts of money spent, and the second argument relates to the possible secular changes in patient satisfaction – that is, what would have happened in the levels of patient satisfaction if the intervention was not implemented in Afghanistan in the first place.

The first argument attempts to provide a glimpse on the dollars' value allocated for the intervention, and the achievements made so far. A total of 12 million dollars was allocated to finance the RBF implementation in the BPHS facilities and EPHS hospitals approximately in one-third of Afghanistan over a three-year period (Afghanistan Ministry of Public Health, , August 2009). Eighty percent of the 12 million dollars was earmarked for paying financial incentives to health facilities and hospitals. According to my study, with the objective of improving patient satisfaction, the current RBF program did not yield significant results within a two-year span. Similarly, RBF did not affect determinants of patient satisfaction at the health facility level. That is, there was so significant improvement in health provider performance, availability of medicine, vaccines, equipment, and physical appearance of health facilities. This suggests that health facilities' capacity to draw more resources (e.g., essential drugs, etc.) from the respective NGO does not appear to have increased as a result of the intervention. Health providers did not improve their communication and interpersonal

skills, and they did not change behaviour to improve quality of care to the satisfaction of patients. Likewise, appearance and cleanliness of health facilities which can be considered as a proxy for health workers' interest in their work environment did not improve over a period of two years. These all suggest that paying monetary incentives to health facilities with the objective of improving quality of care and patient satisfaction may not be properly working if other types of support such as managerial, logistical and non-financial incentives are not provided alongside financial incentives to health workers.

The second argument focuses on the possibility of secular changes in the levels of patient satisfaction over time. The trends in the proportions of patients who were very satisfied with aspects of quality show similar patterns between RBF and non-RBF facilities over time. The trends for both groups of facilities show some improvement in the levels of patient satisfaction on most aspects over time. The reasons for the presence of such trends are not clear. However, one can speculate of at least two possibilities. First, both RBF and non-RBF facilities have been contractually responsible for providing BPHS services, of which the RBF target services form a key component of the BPHS. Further a functional healthcare delivery system has been established in the 11 provinces over the recent ten years. Therefore, even without the RBF program health facilities could have provided the target services. Secondly, due to the publicity of the intervention some sort of spill-over effects may have occurred. This particularly seems to be the case as the rumour spread out that RBF might be scaled up at national level to all health facilities. The publicity of RBF and its possible expansion may have indirectly encouraged non-RBF facilities to compete with RBF facilities so as to secure a better position in the future round of RBF. The above two conditions seem more likely to have played role in driving the secular changes in patient satisfaction in non-RBF facilities.

When it comes to my exploratory question of differentiating RBF effects on patient satisfaction by the three incentive administration types, it was revealed that no clear differences existed between them. This suggests that paying monetary incentives alone – no matter how the incentives are administered in health facilities – may not have the impetus to improve determinants of patient satisfaction at health facility level, and may not improve patient satisfaction over a period of two years.

In this study, I found that income of some types of health providers from RBF increased substantially over time. This was especially the case for those health providers involved directly in the delivery of RBF target services. Percentage of health providers' income from RBF were within the recommended range of 5% to 25% of providers' income commonly practiced in many P4P and performance-based programs worldwide (La Forgia, 2005; Peabody et al., 2011; Young et al., 2007). However, given the seemingly weak purchasing power of Afghani currency in comparison with most countries', the current amounts of financial incentives may not be sufficient to motivate providers for improving performance and quality of care unless other ancillary support and non-financial incentives are also provided to health facilities.

9.2 Policy implications

Considering the apparently weak effects of RBF on some aspects of quality of care and on patient satisfaction, my key message to the donors and policy makers whose goal is to improve the healthcare system is *'do not rely exclusively on paying monetary incentives for healthcare delivery and quality improvement in a post conflict situation'*. In such setting, unless sufficient healthcare resources are supplied and continuous administrative and technical support provided to health providers, the risk of overstretching the scarce resources may compromise the quality of care which can adversely affect patient satisfaction.

The weak effects of RBF on patient satisfaction and on patient satisfaction determinants at health facility level may have to do a lot with the design of the current intervention where except for financial incentives, all managerial and logistical support, including various non-financial incentives, were kept constant between RBF and non-RBF facilities over time. This suggests that relaying solely on paying financial incentives to health providers for quality improvement may not be an effective approach. A more holistic approach focusing on improving various aspects of healthcare system might be a more practical way forward. In future, there are at least six areas the MOPH can capitalize by improving the healthcare system, and may opt to tailor a modified model of the current RBF.

First, the MOPH should focus on tailoring efforts to integrate interventions aiming to improving the healthcare system at various levels. Improving the healthcare system at district level needs more attention as the capacity of health facilities in the district should be

enhanced and maintained so as to strengthen linkages with respective community and various stakeholders present on the ground. This way health facilities may feel more responsibility in terms of meeting patients' and the community's health needs and expectations.

The second area that MOPH can capitalize by tailoring a modified model of RBF relates to the community health workers (CHWs) program. In the current RBF program, the CHWs' role was not explicitly defined – all the efforts were tailored at the health facility level. As the BPHS specifies that CHWs form the backbone of primary healthcare in Afghanistan, they need to be supported. However, this valuable community-based structure has not been fully supported yet. According to a recent study by (Edward et al., 2015) in Afghanistan, it was found that over 75% of CHWs reported unsatisfactory compensation schemes, 69% of them reported inadequate transport, and 40% reported lack of commodities. Our recommendation is in light of the crucial role of CHWs in the delivery of maternal and child health services, particularly in remote and rural areas of Afghanistan. It is imperative to design a RBF scheme for CHWs in the next round of RBF program where both monetary and non-financial incentives need to be considered for them, as research shows that monetary incentives alone cannot be sustained over the long term in many contexts (Glenton et al., 2010).

Third, the critical role of nonfinancial incentives such as supportive supervision, constructive feedback, continuous trainings, better accommodations for staff, and provision of equipment, medicines, and clinical guidelines has already been emphasized when the MOPH issues requests for proposals for bidding NGOs, and when the MOPH signs contracts with them. These nonfinancial incentives need to be strengthened, improved and maintained in the future.

Fourth, instead of paying incentives to health facilities for increasing use of certain target services, the MOPH may opt to resuming the performance bonus practice. With the current design of RBF where health facilities are paid according to the number of extra cases for certain types of services, it seems there is no constructive competition between health facilities within a district. This is because there is no ranking of best performers, nor are there predetermined performance targets for the facilities to achieve. The performance bonus scheme, on the other hand, has already shown some positive results in the PPA contracting approach in Afghanistan.

Fifth, instead of paying performance incentives to health workers, the MOPH may need to revisit the national salary policy, and adjust the salary scales and hardship allowances for health providers if the low motivation of health providers is thought to be due to low payment. This may create some sense of job security and sustainable income for health workers who may focus on delivering the whole range of services specified in the BPHS (or EPHS).

Last, but not least, concerns the MOPH's monitoring and evaluation capacity. The MOPH should focus on strengthening its monitoring capacity as the third party (e.g., Johns Hopkins University) is no longer present in Afghanistan to monitor and evaluate healthcare delivery. In this regard, the MOPH has been assisted by its technical partners over the last decade, and has built the capacity to independently assess the Afghanistan's healthcare system. The national monitoring checklist can be one of the tools which can be used for assessing NGOs' performance against the agreed targets as well as against the standards of the BPHS & EPHS.

9.3 Contributions to the literature on RBF

This study has increased our knowledge about the effects of RBF on patient satisfaction and on its determinants, and about the importance of RBF design and implementation. In my study, I found that paying financial incentives to health facilities neither affects patient satisfaction nor its determinants in the context of a post conflict country. The findings in my study does not support the results reported by several studies conducted in LMICs (Rusa et al., 2009; Huntington et al., 2010; Peabody et al., 2011) because health providers' behaviour did not change, and quality of care and patient satisfaction did not improve as a result of RBF.

My study also highlights the critical role of RBF implementation, particularly when it comes to the administration of incentives. This point is important as no study has been conducted in a post conflict country to investigate whether the way performance incentives are distributed among health workers makes any difference in the quality of care and patient satisfaction. Even if in this study I have not found any difference in the treatment effect between the various ways of incentive administration, it can be a potential area for future research where larger sample sizes with experimental designs can be considered.

In this study, I found that income of various types of health workers from RBF considerably increased over time. Despite this, health provider performance did not improve over a two-year time. This suggests that monetary incentives alone, no matter how they are administered among health workers, may not have the potential to affect quality of care and patient satisfaction in a post conflict setting.

9.4 Methodological considerations

Although this study has increased my knowledge of the treatment effects of financial incentives implemented through four administration types in a fragile state, the findings of this study should be interpreted in light of certain methodological considerations.

One such consideration concerns the drafting of the design for investigating the RBF effects disaggregated by incentive administration types. That is, the selection of administration types was not random and was determined by the managing NGO during program implementation. To create control groups for the four incentive administration types, the non-RBF matched-pair facilities under each type of incentive administration were selected as the control groups. Therefore, the selection process was not exogenous. To address the issue, I use the DID approach which links the treatment effects to change over time for each group and control for the baseline. This can eliminate biases emerging from the unobserved and observed characteristics related to health facilities and patients.

Despite having a relatively sizable sample of health facilities and patients, it is possible that due to the presence of a ceiling effect, I could not detect changes between RBF and non-RBF groups. The ceiling effect seems more plausible when I look into the design of the intervention. That is, health facilities contractually have been responsible for meeting the requirements of the BPHS in terms of availability of health resources, and providing the required primary healthcare for more than a decade in Afghanistan. This suggests that health facilities' capacity has improved over years. Therefore, there was not much room left for RBF to contribute to quality improvement and performance enhancement of health facilities. The problem of ceiling effect may well be the case when it comes to the patient satisfaction measures, because patients often tend to overrate quality of healthcare delivery (Williams, 1994). Therefore, due to the presence of a ceiling effect it is very difficult to detect the difference between the two groups of facilities, unless a large sample is employed.

Another limitation of the study concerns the way data were collected for patient satisfaction. In the questionnaire used in this study, there was no question to identify whether patients used incentivised versus non-incentivised services. Thus, it was not possible for us to stratify treatment effects on the outcomes by the incentivised versus non-incentivised services. If I could estimate RBF treatment effects on the outcomes among users of incentivised services, perhaps the results might have been different from what I have found in this dissertation.

When it comes to external validity and generalization of the findings, there are several issues to be considered. Firstly, the 11 provinces were purposefully selected for the intervention because of the seemingly better accessibility and security condition. These provinces, thus, may not necessarily represent the entire country. However, the findings may provide some indications about RBF effects on some aspects of quality of care, performance of health facilities, and patient satisfaction in a fragile state. Secondly, the study has been conducted in the context where the health system has experienced significant reforms in the last one and half decades, and this has resulted in improved capacity of the MOPH and NGOs in terms of healthcare delivery in the country, especially in the 11 provinces where RBF has been implemented. Therefore, it is possible that I could not detect any difference between RBF and non-RBF due to no obvious difference in their capacities as a result of the intervention.

It is also imperative to note that there may have been other factors such as security concerns in the area that affected facilities' performance, or some additional support provided by other sources to health facilities for which I are not aware of, and I have not unintentionally considered them in the study. In such scenarios it is difficult, if not impossible, to disentangle RBF effects from other factors which either positively or negatively affects the outcome.

9.5 Future research

This study is not an end to the evaluation of RBF, and in its own right it opens up new areas for future research. The crucial role of non-financial incentives to health facilities merits further investigations, as in my study I could not compare between the impact of financial incentives and non-financial incentives on the outcome of interest. In future research, I recommend that effects financial versus non-financial incentives should be investigated. A robust design needs to be used where each type of intervention would have its own control group, and where intervention groups would remain mutually exclusive of each other's effect.

My assumption of the capacity of implementing NGOs in different provinces of Afghanistan was that the organizational capacities of NGOs were similar, as nearly all the NGOs involved in RBF program were Afghan NGOs. I also assumed that the geographic terrains and security situations of all participating provinces were similar. However, they may not necessarily be the case as many dynamic changes can take place in the organization of an NGO during program implementation, or a province may experience several ups and downs in terms of security situation throughout RBF implementation. The above limitations can also be considered potential areas for future research.

At the beginning of this chapter, I highlighted on the total amounts allocated for financing this program. However, due to lack of access to the database on the RBF expenditure I was not able to employ a better analysis to provide an indication on the marginal cost for the target services. In that case, the marginal cost would include all costs incurred for the management, implementation, and evaluation of RBF in Afghanistan. It is imperative to conduct a robust cost effectiveness analysis of the approach, and provide clarity and rationality on the amounts spent so far. The cost effectiveness analysis can be a very interesting topic for the researchers, donors, and policy makers involved in the healthcare delivery in the developing countries.

Likewise, possible unintended consequences of RBF calls upon the interested researchers to undertake this interesting topic. This is particularly relevant in the context of current program where I have identified that generally RBF did not have any impact on patient satisfaction and on some aspects of quality which deal with availability of resources. This might be due to the overstretching of available resources for the use of RBF target services. In my study, I did not have access to the database on the utilization of health services to examine whether RBF has any unintended consequences or whether a positive/negative effect on the utilization of non-incentivized services can be detected.

References

Abadie, A. (2005). Semiparametric difference-in-differences estimators. *The Review of Economic Studies, 72*(1), 1-19.

Abdulhadi, N., Al Shafaee, M., Freudenthal, S., Östenson, C.-G., & Wahlström, R. (2007). Patient-provider interaction from the perspectives of type 2 diabetes patients in Muscat, Oman: a qualitative study. *BMC Health Services Research, 7*(1), 162.

Afghan Public Health Institute (APHI) of the Ministry of Public Health. (2010,). Afghanistan Mortality Survey 2010,.

Afghanistan Ministry of Public Health. (,2005). The Essential Package of Hospital Services for Afghanistan. *Retrieved from http://moph.gov.af/Content/Media/Documents/EPHS-2005-FINAL29122010164126629.pdf*.

Afghanistan Ministry of Public Health. (, August 2009). Operations Manual for Results-based Financing Intervetion in BPHS facilities and hospitals in Afghanistan

Afghanistan Ministry of Public Health. (, December 2011). National Salary Policy for BPHS in Afghanistan.

Afghanistan Ministry of Public Health. (,March 2003). A basic package of health services in Afghanistan. *Retrieved from http://unpan1.un.org/intradoc/groups/public/documents/APCITY/UNPAN018852.pdf*.

Afghanistan. Ministry of Public Health. (2008). Afghansitan health survey 2006: estimates of priority health indicators for rural Afghanistan. *Johns Hopkins University Bloomberg School of Public Health. General Directorate of Policy and Planning, Ministry of Public Health of Afghanistan.*

Aldana, J. M., Piechulek, H., & Al-Sabir, A. (2011). Client satisfaction and quality of care in rural Bangladesh. *Bulletin of the World Health Organization, 76*(6), 512-517.

Ameli, O., & Newbrander, W. (2008). Contracting for health services: effects of utilization and quality on the costs of the Basic Package of Health Services in Afghanistan. *Bulletin of the World Health Organization, 86*(12), 920 - 928.

Andaleeb, S. (2001). Service quality perceptions and patient satisfaction: a study of hospitals in a developing country. *Social science & medicine, 52*, 1359-1370.

Andaleeb, S. (2008). Caring for children: a model of healthcare service quality in Bangladesh. *International Journal for Quality in Health Care, 20*(5), 339-345. doi: 10.1093/intqhc/mzn024

Andaleeb, S., Siddiqui, N., & Khandakar, S. (2007). Patient satisfaction with health services in Bangladesh. *Health policy and planning, 22*(4), 263-273.

Arur, A., Peters, D., Hansen, P., Mashkoor, M. A., Steinhardt, L. C., & Burnham, G. (2010). Contracting for health and curative care use in Afghanistan between 2004 and 2005. *Health policy and planning, 25*(2), 135-144.

Ashrafun, L., & Uddin, M. J. (2011). Factors Determining Inpatient Satisfaction with Hospital Care in Bangladesh. *Asian Social Science, 7*(6), 15-24.

Atinga, R. A., Abekah-Nkrumah, G., & Domfeh, K. A. (2011). Managing healthcare quality in Ghana: a necessity of patient satisfaction. *Int J Health Care Qual Assur, 24*(7), 548-563.

Babakus, E., & Boller, G. W. (1992). An Empirical Assessment of the SERVQUAL Scale. *Journal of Business Research, 24*, 253-268.

Baltussen, R., Ye, Y., Haddad, S., & Sauerborn, R. (2002). Perceived quality of care of primary health care services in Burkino Faso. *Health policy and planning, 17*(1), 42-48.

Bartlett, L. A., Mawji, S., Whitehead, S., Crouse, C., Dalil, S., Ionete, D., . . . Team, A. M. M. S. (2005). Where giving birth is a forecast of death: maternal mortality in four districts of Afghanistan, 1999–2002. *The Lancet, 365*(9462), 864-870.

Basinga, P., Gertler, P. J., Binagwaho, A., Soucat, A. L. B., Sturdy, J., & J., V. C. M. (2011). Effect on maternal and child health services in Rwanda of payment to primary health-care providers for performance: an impact evaluation. *Lancet, 377*(9775), 1421-1428.

Batley, R., & Mcloughlin, C. (2010). Engagement with Non-State Service Providers in Fragile States: Reconciling State-Building and Service Delivery. *Development Policy Review, 28*(2), 131-154.

Bernhart, M. H., Waidnyana, I. G. P., Wihardjo, H., & Pohan, I. (1999). Patient satisfaction in developing countries. *Social science & medicine, 48*, 989-996.

Bhushan, I., Bloom, E., Clingingsmith, D., Hong, R., King, E., Kremer, M., . . . Schwartz, J. B. (2007). Contracting for Health: Evidence from Cambodia. 1 - 47.

Bhushan, I., Keller, S., & Schwartz, B. (2002). Achieving the Twin Objectives of Efficiency and Equity: Contracting Health Services in Cambodia. *ERD Policy Brief No.6*.

Bhutta, Z., Gupta, I., de'Silva, H., Manandhar, D., Awasthi, S., Hossain, S. M. M., & Salam, M. A. (2004). Maternal and child health: is South Asia ready for change? *British Medical Journal, 328*, 816-819.

Bitrán, R., Má, C., & Gómez, P. (2005). The San Miguelito Hospital Reform in Panama-Evaluation and Lessons. *WORLD BANK WORKING PAPER, 57*, 89.

Bokhour, B. G., Burgess, J. F., Hook, J. M., White, B., Berlowitz, D., Guldin, M. R., . . . Young, G. J. (2006). Incentive implementation in physician practices: a qualitative study of practice executive perspectives on pay for performance. *Medical Care Research and Review, 63*(1 suppl), 73S-95S.

Bower, H. (2002). Reconstructing Afghanistan's health system: are lessons being learned from the past. *Masters of Science Project. LSHTM: London*.

Bowser, D., Figueroa, R., Natiq, L., & Okunogbe, A. (2013). A preliminary assessment of financial stability, efficiency, health systems and health outcomes using performance-based contracts in Belize. *Global Public Health*.

Bristol, N. (2005). Reconstructing Afghanistan's health system. *Lancet, 366*, 2075-2076.

Campbell, S. M., Reeves, D., Kontopantelis, E., Sibbald, B., & Roland, M. (2009). Effects of Pay for Performance on the Quality of Primary Care in England. *New England Journal of Medicine, 361*(4), 368-378.

Cardozo, R. N. (1965). An experimental study of customer effort, expectation, and satisfaction. *Journal of marketing research*, 244-249.

Carman, J. (1990). Consumer Perceptions of Service Quality: An Assessment of the SERVQUAL Dimensions. *Journal of Retailing, 66*(1).

Carmine, E., & Zeller, R. (1979). Reliability and validity assessment. *SAGE Publications/ Beverly Hills/London*.

Carthaigh, N. N., De Gryse, B., Esmati, A. S., Nizar, B., Van Overloop, C., Fricke, R., . . . Philips, M. (2014). Patients struggle to access effective health care due to ongoing violence, distance, costs and health service performance in Afghanistan. *International health*, ihu086.

Cercone, J., Briceño, R., & Gauri, V. (2005). Contracting Primary Health Care Services-The Case of Costa Rica. *WORLD BANK WORKING PAPER, 57*, 109.

Cockcroft, A., Khan, A., Ansari, N. M., Omer, K., Hamel, C., & Andersson, N. (2011). Does contracting of health care in Afghanistan work? Public and service-users' perceptions and experience. *BMC Health Services Research, 11*(Suppl 2), S11.

Cronin, J., & Taylor., S. (1992). Measuring Service Quality: A Reexamination and Extension. *Journal of Marketing, 56*, 55-68.

Custers, T., Hurley, J., Klazinga, N. S., & Brown, A. D. (2008). Selecting effective incentive structures in health care: A decision framework to support health care purchasers in finding the right incentives to drive performance. *BMC Health Services Research, 8*(1), 66.

Dalil, S., Newbrander, W., Loevinsohn, B., Naeem, A. J., Griffin, J., Salama, P., & Momand, F. M. (2014). Aid effectiveness in rebuilding the Afghan health system: A reflection. *Global Public Health, 9*(sup1), S124-S136.

Danel, I., & La Forgia, G. (2005). Contracting for basic health care in rural Guatemala: comparison of the performance of three delivery models. *Health Systems Innovations in Central America: Lessons and Impact of New Approaches*, 49-88.

De Geyndt, W. (1995). *Managing the quality of health care in developing countries* (Vol. 258): World Bank Publications.

Donabedian, A. (1988). The assessment of technology and quality: A comparative study of certainties and ambiguities. *International Journal of Technology Assessment in Health Care, 4*(04), 487-496.

Edward, A., Branchini, C., Aitken, I., Roach, M., Osei-Bonsu, K., & Arwal, S. H. (2015). Toward universal coverage in Afghanistan: A multi-stakeholder assessment of capacity investments in the community health worker system. *Social science & medicine*.

Edward, A., Dwivedi, V., Mustafa, L., Hansen, P. M., Peters, D. H., & Burnham, G. (2009). Trends in the quality of health care for children aged less than 5 years in Afghanistan, 2004-2006. *Bulletin of the World Health Organization, 87*, 940-949.

Edward, A., Kumar, B., Kakar, F., Salehi, A. S., Burnham, G., & Peters, D. H. (2011). Configuring Balanced Scorecards for Measuring Health System Performance: Evidence from 5 Year's Evaluation in Afghanistan. *PLOS Medicine, 8*(7), 1-9.

Eichler, R. (2006). Can "Pay for Performance" Increase Utilization by the Poor and Improve the Quality of Health Services?Discussion paper for the first meeting of the Working Group on Performance-Based Incentives. *Center for Global Development*.

Eichler, R., Auxilla, P., Antoine, U., & Desmangles, B. (2007). Performance-Based Incentives for Health: Six Years of Results from Supply-Side Programs in Haiti. *CGC Working Paper # 121, Center for Global Development*.

Eldridge, C., & Palmer, N. (2009). Performance-based payment: some reflections on the discourse, evidence and unanswered questions. *Health policy and planning, 24*(3), 160-166.

Evans, D. B., Edejer, T. T. T., Lauer, J., Frenk, J., & Murray, C. J. (2001). Measuring quality: from the system to the provider. *International Journal for Quality in Health Care, 13*(6), 439-446.

Field Office of Johns Hopkins in Kabul Afghanistan. (, November 2010). Verification report - Monitoring and Evaluation of Results-based Financing

Franco, L. M., Bennett, S., & Kanfer, R. (2002). Health sector reform and public sector health worker motivation: a conceptual framework. *Social science & medicine, 54*(8), 1255-1266.

Franco, L. M., Bennett, S., Kanfer, R., & Stubblebine, P. (2004). Determinants and consequences of health worker motivation in hospitals in Jordan and Georgia. *Social science & medicine, 58*(2), 343-355.

Gallouj, C. (1997). Asymmetry of information and the service relationship: selection and evaluation of the service provider. *International Journal of Service Industry Management, 8*(1), 42-64.

Giustozzi, A. (2009). Afghanistan: getting worse before getting better?

Glenton, C., Scheel, I. B., Pradhan, S., Lewin, S., Hodgins, S., & Shrestha, V. (2010). The female community health volunteer programme in Nepal: decision makers' perceptions of volunteerism, payment and other incentives. *Social science & medicine, 70*(12), 1920-1927.

Gordis. Leon. (2009). Epidemiological Methods. *4th ed*, 103-105.

Gorter, A., Ir, P., & Meessen, B. (2013). Evidence review: Results-based financing of maternal and newborn health care in low-and lower-middle-income countries. *study commissioned and funded by the German Federal Ministry for Economic Cooperation and Development through the sector project PROFILE at Deutsche Gesellschaft für Internationale Zusammenarbeit, Bonn & Eschborn, Germany*.

Gustavo Sanchez, S. (June 2013). Bootstrap with panel data. *retrieved from: http://www.stata.com/support/faqs/statistics/bootstrap-with-panel-data/*.

Haddad, S. S., Fournier, P. P., Machouf, N. N., & Yatara, F. F. (1998). What does quality mean to lay people? Community perceptions of primary health care services in Guinea. *Social science & medicine (1982), 47*(3), 381-394.

Hansen, Peters, D. H., Viswanathan, K., Rao, K. D., Mashkoor, A., & Burnham, G. (2008). Client perception of the quality of primary care services in Afghanistan. *International Journal for Quality in Health Care, 20*(6), 384-391.

Hansen, P. M., Peters, D. H., Edward, A., Gupta, S., Arur, A., Niayesh, H., & Burnham, G. (2008). Determinants of primary care service quality in Afghanistan. *International Journal for Quality in Health Care, 20*(6), 375-383.

Hansen, P. M., Peters, D. H., Niayesh, H., Singh, L. P., Dwivedi, V., & Burnham, G. (2008). Measuring and managing progress in the establishment of basic health services: the Afghanistan Health Sector Balanced Scorecard. *International Journal of Health Planning and Management, 23*, 107-117.

Harman, H. H. (1976). Modern Factor Anaysis. *The University of Chicago, Press. Ltd. London*.

Hecht, R., Batson, A., & Brenzel, L. (2004). Making Health Care Accountable, Why performance-based financing of health services in developing countries is getting more attention. *Finance and Management March 2004*.

Henderson, L. N., & Tulloch, J. (2008). Incentives for retaining and motivating health workers in Pacific and Asian countries. *Human resources for health, 6*(1), 18.

Hogan, M. C., Foreman, K. J., Naghavi, M., Ahn, S. Y., Wang, M., Makela, S. M., . . . Lozano, R. (2010). Maternal mortality for 181 countries, 1980-2008: a systemic analysis of progress towards Millenium Development Goal 5. *Lancet, 375*, 1609-1623.

Huntington, D., Zaky, H. H., Shawky, S., Fattah, F. A., & El-Hadary, E. (2010). Impact of a service provider incentive payment scheme on quality of reproductive and child-health services in Egypt. *J Health Popul Nutr, 28*(3), 273.

Ireland, M., Paul, E., & Dujardin, B. (2011). Can performance-based financing be used to reform health systems in developing countries? *Bulletin of the World Health Organization, 89*, 695-698.

Islamic Republic of Afghanistan. Afghanistan National Development Strategy (2008-2013). *Retrieved from http://www.undp.org.af/Publications/KeyDocuments/ANDS_Full_Eng.pdf*.

Kirschner, K., Braspenning, J., Jacobs, A., & Grol, R. (2012). Design choices made by target users for a pay-for-performance program in primary care: an action research approach. *BMC Fam Pract, 13*(25), 1-10.

La Forgia, G. M. (2005). *Health system innovations in Central America: lessons and impact of new approaches*: World Bank Publications.

Lindahl, A. K., Oxman, A., & Fretheim, A. (2008). An overview of research on the effects of results-based financing. *Norwegian Knowledge Centre for the Health Services*.

Liu, X., Hotchkiss, D. R., & Bose, S. (2008). The effectiveness of contracting-out primary health care services in developing countries: a review of the evidence. *Health policy and planning, 23*(1), 1-13.

Loevinsohn, B., Couffinhal, A., & Pande, A. (2009). Contracting-in management to strengthen publicly financed primary health services—The experience of Punjab, Pakistan. *Health Policy, 91*(1), 17-23.

Loevinsohn, B., & Harding, A. (2005). Buying results? Contracting for healht service delivery in developing countries. *Lancet, 366*, 676-681.

Manongi, R. N., Marchant, T. C., & Bygbjerg, I. C. (2006). Improving motivation among primary health care workers in Tanzania: a health worker perspective. *Human resources for health, 4*(1), 6.

Mathauer, I., & Imhoff, I. (2006). Health worker motivation in Africa: the role of non-financial incentives and human resource management tools. *Human resources for health, 4*(1), 24.

Mbindyo, P., Gilson, L., Blaauw, D., & English, M. (2009). Contextual influences on health worker motivation in district hospitals in Kenya. *Implementation Science, 4*(1), 43.

Meessen, B., Kashala, J.-P., & Musango, L. (2007). Output-based payment to boost staff productivity in public health centres: contracting in Kabutare district, Rwanda. *Bulletin of the World Health Organization, 85*(2), 108-115.

Meessen, B., Musango, L., Kashala, J.-P., & Lemlin, J. (2006). Reviewing institutions of rural health centres: the Performance Initiative in Butare, Rwanda. *Tropical Medicine and International Health, 11*(8), 1303-1317.

Meessen, B., Soucat, A., & Sekabarga, C. (2011). Performance-based financing: just a donor fad or a catalyst towards comprehensive health-care reform? *Bulletin of the World Health Organization, 89*, 153-156.

Meyer, B. D. (1995). Natural and quasi-experiments in economics. *Journal of business & economic statistics, 13*(2), 151-161.

Michael, M., Pavignani, E., & Hill, P. S. (2013). Too good to be true? An assessment of health system progress in Afghanistan, 2002–2012. *Medicine, Conflict and Survival, 29*(4), 322-345.

Minami, N. A. (2010). Understanding the Problem in Afghanistan A Plan for Stability.

Morgan, L. (2010). Some days are better than others: lessons learned from Uganda's first results-based financing pilot. *The World Bank*.

Mostafa, M. M. (2005). An empirical study of patients' expectations and satisfactions in Egyptian hospitals. *Int J Health Care Qual Assur, 18*(7), 516-532.

Newbrander, W., Ickx, P., Feroz, F., & Stanekzai, H. (2014). Afghanistan's Basic Package of Health Services: Its development and effects on rebuilding the health system. *Global Public Health*(ahead-of-print), 1-23.

Newbrander, W., Yoder, R., & Debevoise, A. B. (2007). Rebuilding health systems in post-conflict countries: estimating the costs of basic services. *International Journal of Health Planning and Management, 22*, 319-336.

Newsome, P., & Wright, G. H. (1999). Patient Management: A review of patient satisfaction: 1. Concepts of satisfaction. *British dental journal, 186*(4), 161-165.

Oliver, R. L. (1980). A cognitive model of the antecedents and consequences of satisfaction decisions. *Journal of marketing research*, 460-469.

Olken, B. A., Onishi, J., & Wong, S. (2012). Should Aid Reward Performance? Evidence from a field experiment on health and education in Indonesia: National Bureau of Economic Research.

Oxman, A. D., & Fretheim, A. (2009). Can paying for results help to achieve the Millennium Development Goals? Overview of the effectiveness of results-based financing. *Journal of Evidence-Based Medicine, 2*(2), 70-83.

Palmer, N. N., Strong, L. L., Wali, A. A., & Sondorp, E. E. (2006). Contracting out health services in fragile states. *BMJ (Clinical research ed.), 332*(7543), 718-721.

Parasuraman, A., Berry, L., & Zeithaml, V. (1988). SERQUAL: A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality. *Journal of Retailing, 64*(1).

Parasuraman, A., Berry, L., & Zeithaml, V. (1991). Refinement and Reassessment of the SERVQUAL Scale. *Journal of Retailing, 67*(4), 420 - 450.

Peabody, J., Shimkhada, R., Quimbo, S., Florentino, J., Bacate, M., McCulloch, C. E., & Solon, O. (2011). Financial Incentives And Measurement Improved Physicians' Quality Of Care In The Philippines. *Health Affairs, 30*(4), 773-781.

Penneys, N. S. (1997). Quality: its definition, measurement, and applications in dermatology. *J Am Acad Dermatol, 37*(3), 503-507.

Peters, D. H., Noor, A. A., Singh, L. P., Kakar, F. K., Hansen, P. M., & Burnham, G. (2007). A balanced scorecard for health services in Afghanistan. *Bulletin of the World Health Organization, 85*, 146-151.

Petersen, L. A., Woodard, L. D., Urech, T., Daw, C., & Sookanan, S. (2006). Does pay-for-performance improve the quality of health care? *Annals of internal medicine, 145*(4), 265-272.

Peterson, R. A., & Wilson, W. R. (1992). Measuring customer satisfaction: fact and artifact. *Journal of the Academy of Marketing science, 20*(1), 61-71.

Press, I., & Fullam, F. (2011). Patient satisfaction in pay for performance programs. *Qual Manag Health Care, 20*(2), 110-115. doi: 10.1097/QMH.0b013e318213aed0

Rao, K. D. K. D., Peters, D. H. D. H., & Bandeen-Roche, K. K. (2006). Towards patient-centered health services in India--a scale to measure patient perceptions of quality. *Int J Qual Health Care, 18*(6), 414-421.

Rasooly, M. H., Govindasamy, P., Aqil, A., Rutstein, S., Arnold, F., Noormal, B., . . . Shadoul, A. (2014). Success in reducing maternal and child mortality in Afghanistan. *Global Public Health, 9*(sup1), S29-S42.

Reerink, I. H., & Sauerborn, R. (1996). Quality of primary health care in developing countries: recent experiences and future directions. *International Journal for Quality in Health Care, 8*(2), 131-139.

Ridde, V. (2005). Performance-based partnership agreements for the reconstruction of the health system in Afghanistan. *Development in Practice, 15*(1), 4-15.

Robyn, P. J., Bärnighausen, T., Souares, A., Savadogo, G., Bicaba, B., Sié, A., & Sauerborn, R. (2012). Health worker preferences for community-based health insurance payment mechanisms: a discrete choice experiment. *BMC Health Services Research, 12*(1), 159.

Rosenthal, M. B., & Frank, R. G. (2006). What is the empirical basis for paying for quality in health care? *Medical Care Research and Review, 63*(2), 135-157.

Roski, J., Jeddeloh, R., An, L., Lando, H., Hannan, P., Hall, C., & Zhu, S.-H. (2003). The impact of financial incentives and a patient registry on preventive care quality: increasing provider adherence to evidence-based smoking cessation practice guidelines. *Preventive medicine, 36*(3), 291-299.

Rusa, L., Ngirabega, J. d. D., Janssen, W., Van Bastelaere, S., Porignon, D., & Vandenbulcke, W. (2009). Performance-based financing for better quality of services in Rwandan health centres: 3-year experience. *Tropical Medicine & International Health, 14*(7), 830-837.

Sabri, B., Siddiqui, S., Ahmed, A., Kakar, F., & Perrot, J. (2007). Towards sustainable delivery of health services in Afghanistan: options for the future. *Bulletin of the World Health Organization, 85*(9), 712-719.

Safavi, K. (2006). Patient-Centered Pay for Performance: Are We Missing the Target? *Journal of Healthcare Management, 51*(4), 215-218.

Sekabarga, C., Diop, F., & Soucat, A. (2011). Can innovative health financing policies increase access to MDG-related services? Evidence from Rwanda. *Health policy and planning, 26*. doi: 10.1093/heapol/czr070

Shen, Y. (2003). Selection incentives in a performance-based contracting system. *Health Serv Res, 38*(2), 535-552.

Siddiqi, S., Masud, T. I., & Sabri, B. (2006). Contracting but not without caution: experience with outsourcing of health services in countries of the Eastern Mediterranean Region. *Bulletin of the World Health Organization, 84*(11), 867-875.

Silk Route Training & Research Organization. (2011). Parliamentartians' expectations and perceptions of the Afghanistan health sector. *Not published*.

Smith, P. C., & York, N. (2004). Quality incentives: the case of UK general practitioners. *Health Affairs, 23*(3), 112-118.

Soeters, R., & Griffiths, F. (2003). Improving government health services through contract management: a case from Cambodia. *Health policy and planning, 18*(1), 74-83.

Soeters, R., Habineza, C., & Peerenboom, P. B. (2006). Performance-based financing and changing the district health system: experience from Rwanda. *Bulletin of the World Health Organization, 84*(11), 884-889.

Soeters, R., Peerenboom, P. B., Mushagalusa, P., & Kimanuka, C. (2011). Performance-Based Financing Experiment Improved Health Care in The Democratic Republic of Congo. *Health Affairs, 30*(8), 1518-1527.

Soeung, S., Grundy, J., Ly, C., Samnang, C., Boreland, M., Maynard, J., & Biggs, B. (2006). Improving immunization coverage through budgeted microplans and sub-national performance

agreements: early experience from Cambodia. *Asia Pacific Journal of Public Health, 18*(1), 29-38.
Sondorp, E. (2004). A time-seried analysis of health service delivery in Afghanistan. *Service Delivery in Difficult Environment, DFID Health Systems Resource Centre, 2004*.
Strong, L., Wali, A., & Sondorp, E. (2005). Health Policy in Afghanistan: two years of rapid change A review of the process from 2001 to 2003. *London School of Hygiene & Tropical Medicine*.
The World Bank. (2003). Private sector assessment for health, nutrition and population in Bangladesh.
Town, R., Wholey, D. R., Kralewski, J., & Dowd, B. (2004). Assessing the influence of incentives on physicians and medical groups. *Medical Care Research and Review, 61*(3 suppl), 80S-118S.
Travis, P., Bennett, S., Haines, A., Pang, T., Bhutta, Z., Huder, A. A., . . . Evans, T. (2004). Overcoming health-systems constraints to achieve the Millennium Development Goals. *Lancet, 364*, 900-906.
Tynkkynen, L.-K., Keskimäki, I., & Lehto, J. (2013). Purchaser–provider splits in health care—the case of Finland. *Health Policy, 111*(3), 221-225.
Van Herck, P., De Smedt, D., Annemans, L., Remmen, R., Rosenthal, M. B., & Sermeus, W. (2010). Systematic review: effects, design choices, and context of pay-for-performance in health care. *BMC Health Services Research, 10*(1), 247.
Vuori, H. (1987). Patient satisfaction--an attribute or indicator of the quality of care? *QRB. Quality review bulletin, 13*(3), 106-108.
Waldman, R., & Newbrander, W. (2014). *Afghanistan's health system: Moving forward in challenging circumstances 2002–2013*: Global Public Health: An International Journal for Research, Policy and Practice.
Waldman, R., Strong, L., & Wali, A. (2006). Afghanistan's Health System Since 2001: Condition Improved, Prognosis Cautiously Optimistic. *Afghanistan Research and Evaluation Unit (AREU), Briefing Paper Series*, 1-22.
Williams, B. (1994). Patient satisfaction: a valid concept? *Social science & medicine, 38*(4), 509-516.
Willis-Shattuck, M., Bidwell, P., Thomas, S., Wyness, L., Blaauw, D., & Ditlopo, P. (2008). Motivation and retention of health workers in developing countries: a systematic review. *BMC Health Services Research, 8*(1), 247.
Witter, S., Fretheim, A., Kessy, F. L., & Lindahl, A. K. (2012). Paying for performance to improve the delivery of health interventions in low- and middle-income countries. *Cochrane database of systematic reviews (Online), 2*, CD007899-CD007899.
Young, G., Douglas A, C., & Andrew, F. (2007). Practical Issues in the Design and Implementation of Pay-for-Quality Programs. *Journal of Healthcare Management, 52*(1), 10-19.