The Effects of Health Sector Reform Interventions in Egypt on Family Planning and Maternal and Child Health

Amira El-Shal

A thesis submitted in part-fulfilment of the requirements for the degree of
Doctor of Philosophy in Economics
City, University of London
Department of Economics

February, 2017

Contents

List of Tables

List of Figures

Acknowledgments

First of all, all praises are due to Allah, the Almighty, for guiding me along the path. Without His blessings, this achievement would not have been possible.

I would like to thank my loving parents who have supported and taken care of me to get to this point in my life and career. I am privileged to have the best parents, sister and brother one could have. My utmost regard goes to my mother, Heba Gamal El-Din, without whom I would not have been who I am today. I would also like to thank my grandparents for being a constant source of inspiration.

My sincere gratitude goes to my distinguished supervisor and mentor, Professor Mireia Jofre-Bonet, for her invaluable support and consistent motivation. She has always made herself available to clarify my doubts, even when I was still applying to City, despite her very busy schedule. It has been a great opportunity to pursue my Ph.D. under her guidance and to learn from her research expertise.

My sincere gratitude extends to my second supervisor, Dr. Patricia Cubi-Molla, for our constructive discussions, for her insightful comments and feedback, and for her kindness especially in tough times.

Finally, I wish to express my profound appreciation to my dearest friend and companion, Eman Moustafa, who has always been with me through the peaks and valleys. As I count my blessings, I consider myself tremendously fortunate to have someone with whom I share moments of deep anxiety but also of big excitement.

Abstract

This thesis is composed of four essays that make empirical contributions to impact evaluations of health sector interventions in low- and middle-income countries, in light of the interventions introduced under Egypt's Health Sector Reform Program (HSRP) between 2000 and 2014. We are mainly interested in the effects on family planning and maternal and child health. Different methods are used in this context: difference-in-differences (DD), DD propensity score matching (PSM), fixed effects (FE), random effects (RE) and pooled ordinary least-squares (POLS).

In chapter 3, we estimate the effect of improving the quality of health care through facility accreditation on the family planning, maternal health and child health outcomes that we expect to reflect the effect of compliance with quality standards, policies and procedures. We found that accreditation had multiple positive effects, especially on delivery care and child morbidity prevalence. No significant effects were observed, however, with respect to most antenatal care (ANC) outcomes.

In chapter 4, we estimate the medium-term effect of introducing user fees on the utilization of family planning, ANC and delivery care services, women's access to health care, and child health status. With respect to ANC, we found that the positive effect of increased willingness to pay for an improved quality of service outweighed the negative effect of the price elasticity of demand. Introducing user fees was associated with a higher likelihood of receiving ANC by skilled health personnel, a higher likelihood of receiving at least four ANC visits and a higher likelihood of receiving iron supplements during pregnancy. However, the two effects offset each other with respect to the outcomes that reflect the utilization of family planning and delivery care services, women's access to health care, and child health status. No net effect at all was observed on these outcomes.

Chapter 5 complements the analysis of chapter 4 by allowing us to estimate the net effect of combining user fees and two quality improvement interventions: facility accreditation and performance-based financing (PBF). Again, we observe positive effects on both the utilization and the quality of ANC services. More notably, a positive effect on access to care was observed during our first study period that is more likely to reflect the effect of quality improvements. These effects, however, were reversed during the second study period that is more likely to reflect the effect of user fee introduction. The positive effects reported in chapters 4 and 5 were mainly with respect to ANC. No effects were reported on the outcomes that reflect the utilization of family planning and delivery care services, and child health status.

In chapter 6, we estimate the effect of discontinuing provider incentives on health outcomes that reflect the health services targeted by the PBF scheme as well as the quality of these services. We found that discontinuing the incentives had a negative effect on four out of seven health outcomes: knowledge of contraceptive methods, receiving ANC by skilled health personnel, receiving iron supplements during pregnancy and, more importantly, under-five child mortality.

Our findings, first, suggest that improving the quality of care through facility accreditation could be particularly effective in improving delivery care and child health. However, a high level of commitment from the central government is indispensable to sustain the positive effects of quality improvement interventions. Second, introducing user fees will not necessarily have negative effects on access to and utilization of family planning, maternal health and child health services. However, user fees are ineffective, in general, as a stand-alone policy. Third, negative effects of introducing user fees in low- and middle-income settings on the utilization of healthcare services can be mitigated by officially exempting the poor from any fees at the point of service. More importantly, this exemption should be known to the population. Fourth, combining quality improvement interventions with user fees will not necessarily add to the few positive effects obtained when user fees are introduced as a stand-alone policy. Finally, provider incentives should be introduced carefully in low- and middle-income countries as negative effects are observed when these incentives are discontinued.

List of Abbreviations

ACE-I	Angiotensin-converting enzyme inhibitor
AMI	Absolute maternal indications
ANC	Antenatal care
ARB	Angiotensin receptor blocker
ARI	Acute respiratory infection
BBP	Basic benefit package
BP	Blood pressure
C-section	Cesarean section
CAPMAS	Central Agency for Public Mobilization and Statistics
DD	Difference-in-differences
DHS	Demographic and Health Survey
EGP	Egyptian pound
EMW	Ever-married women
FE	Fixed effects
FHC	Family Health Center
FHF	Family Health Fund
FHM	Family Health Model
GLS	Generalized least squares
GOE	Government of Egypt
GPS	Global Positioning System
HH	Household
HIO	Health Insurance Organization
HSRP	Health Sector Reform Program
IEC	Information/education/communication
IMCI	Integrated Management of Child Illnesses
ITN	Insecticide-treated bed net
IUD	Intrauterine device
IV	Instrumental variables
LM	Lagrangian multiplier
MDG	Millennium Development Goal
MENA	Middle East and North Africa
MIS	Management Information System
MOH	Ministry of Health
NICHP	National Information Center for Health and Population
P4P	Pay-for-performance
PBF	Performance-based financing
PHC	Primary health care
POLS	Pooled ordinary least-squares
Ppts	Percentage points
PSM	Propensity score matching
QIP	Quality Improvement and Accreditation Program
QRU	Quality and Regulation Unit
RE	Random effects
TB	Tuberculosis
UN	United Nations
VA	Veterans Health Administration
WHO	World Health Organization

1. INTRODUCTION

1.1 MOTIVATION

Improving the health and well-being of mothers and children is an ultimate public health goal for all countries. There has been progress in maternal and child health indicators in many low- and middle-income countries. However, progress was not enough to achieve the health-related Millennium Development Goals (MDGs) adopted by the United Nations (UN) in 2000. Moreover, the progress achieved was uneven, with the poorer countries lagging behind the rest.

Contraceptive use remains to be low in the least developed countries (40 percent) and is particularly low in Africa (33 percent) (United Nations, 2015). Antenatal care (ANC) coverage (at least four visits) was as low as 54 percent in Africa and 56 percent in the Eastern Mediterranean in 2013. Only 54 percent and 59 percent of women in Africa and South-East Asia, respectively, benefited in 2013 from skilled care during childbirth. A total of 303,000 women died of complications during pregnancy or childbirth in 2015, with 64 percent of these deaths occurring in Africa only. In 2015, an estimated 5.9 million children worldwide died before reaching the age of five. The heaviest burden of under-five mortality was borne by Africa (2.8 million) followed by South-East Asia (1.6 million) (World Health Organization, 2016).

The progress achieved with respect to maternal and child health was triggered by different health sector interventions across countries. In 1997, Egypt introduced the Health Sector Reform Program (HSRP) to address persistent needs in maternal and child health through an emphasis on primary health care (PHC). Different types of interventions were introduced simultaneously. Three supply-side interventions were introduced to strengthen health service delivery: renewal of PHC infrastructure and equipment, development of human resource and quality assurance. In parallel, two interventions were introduced to improve healthcare financing: performance-based financing (PBF) on the supply side and user fees on the demand side.

In this context, this thesis investigates the effects of different interventions under Egypt's HSRP on family planning and maternal and child health between 2000 and 2014. We provide reliable empirical evidence on these interventions for low- and middle-income countries in general and for Middle East and North Africa (MENA) in particular.

1.2 RESEARCH OBJECTIVES

This thesis has the five main objectives. The first objective is to estimate the effect of improving the quality of health care through facility accreditation on family planning, maternal health and child health outcomes during the period 2000-2008. The second is to estimate the effect of introducing user fees on family planning, maternal health and child health outcomes during the period 2008-2014. Our third objective is to estimate the effect of introducing user fees together with quality improvement interventions on family planning, maternal health and child health outcomes during the period 2000-2008. The fourth objective is to estimate the effect of discontinuing provider incentives on family planning, maternal health and child health outcomes during the period 2005-2014. The final objective of this thesis is to investigate complementary methods and datasets in assessing the effectiveness of stand-alone as well as simultaneous interventions under Egypt's HSRP.

1.3 MAIN CONTRIBUTION

This thesis makes four important contributions. First, it provides a comprehensive framework for investigating the effects of different health sector interventions based on differential reform phases. In this context, robust methods and multiple datasets are used complementarily to ensure the reliability of the results.

Second, this thesis disaggregates the effects of different types of supply- and demand-side interventions that were introduced simultaneously. Such disaggregation allows us to compare the effectiveness of the different types of interventions. Moreover, we compare between the effects of interventions that were introduced simultaneously and the effects of stand-alone interventions.

Third, the thesis investigates the effects of interventions on a complete set of health outcomes that cover not only the quantity, but also the quality, of the healthcare services provided under the HSRP. Whenever applicable, outcome measures of

quality of healthcare services are included in the analyses together with the conventional measures of the quantity of services received.

Fourth, there is several key policy implications that could be drawn from the results of this thesis. For example, our findings suggest that facility accreditation alone may not be sufficient to sustain high quality of health care in low- and middle-income settings. Our findings also suggest that, in general, user fees are ineffective as a stand-alone policy. However, we found evidence that, even when accompanied by quality improvements, introducing user fees in low- and middle-income settings can have negative effects on access to and utilization of health care. Our findings as well suggest that PBF schemes need to be applied carefully in low- and middle-income countries.

1.4 THESIS OUTLINE

In this chapter, we discuss the motivation of the thesis, outline its main research objectives and highlight its key contribution to the current knowledge base. Chapter 2 describes the context in which Egypt's HSRP has been introduced and the data sources used to investigate the effects of the program.

In chapter 3, we estimate the effects of quality improvement through a facility accreditation program between 2000 and 2008 on the health outcomes that we expect to reflect the effect of compliance with quality standards, policies and procedures, which are the focus of accreditation assessment. Chapter 4 provides evidence on the medium-term effect of introducing user fees in a middle-income setting on the utilization of family planning, ANC and delivery care services, women's access to health care, and child health status during the period 2008-2014. Chapter 5 complements our analysis in chapter 4 by estimating the net effect of combining user fees and quality improvement interventions during the period 2000-2008 instead of estimating the effect of introducing user fees as a stand-alone policy. In chapter 6, we estimate the effects of discontinuing provider incentives in Egypt, after being in operation for more than five years, on family planning, maternal health and child health outcomes that reflect the health services targeted by the PBF scheme as well as the quality of these services in contracted facilities.

Chapter 7 summarizes the main conclusions of the thesis and recommends some directions for future research.

2. BACKGROUND AND DATA SOURCES

2.1 INTRODUCTION

In this chapter, we describe the context in which Egypt's Health Sector Reform Program (HSRP) was introduced and present the sources of the data used to investigate the effects of the program.

First, we discuss the systemic problems in Egypt's healthcare system, including poor health outcomes, inequity, inaccessibility, inefficiency, poor quality and long-run financial instability. These along with other problems helped trigger healthcare reform. We also describe the integrated package of service delivery and financing interventions introduced under the HSRP. We indicate the progress of the HSRP implementation as well.

Second, we discuss different sources of data our study draws on, including facility-level data, district-level data, demographic and health surveys (DHS) as well as other supplemental sources of information.

With this in mind, the remainder of the chapter is structured as follows: section 2.2 gives an overview of the HSRP; section 2.3 describes the sources of our data; and section 2.4 concludes.

2.2 Egypt's HSRP

In this section, we discuss the challenges that triggered the HSRP, highlight its objective and guiding principles, and describe the two main components of the program and the interventions introduced under each component.

The providers of primary health care (PHC) in Egypt are the Ministry of Health (MOH) (public), the Health Insurance Organization (HIO) (public) and the private sector. MOH operates a nationwide network of PHC facilities. These public PHC facilities serve as the "insurer of last resort" by offering free or substantially subsidized PHC services to uninsured individuals.

In the early 1990s, the main focus of Egypt's national health strategy was to achieve universal health coverage. The Government of Egypt (GOE) pressured HIO - Egypt's largest health insurer[1]- to rapidly expand coverage to new groups, such as infants and school children. The pressing need to achieve rapid progress towards universal coverage did not take into account the appropriate health system financing required to ensure sustainability. The quality of healthcare services was disregarded as well.

The HSRP was triggered by several challenges. Egypt's health outcomes were unacceptably poor, and even worse than the lower middle-income average, especially maternal and child outcomes. The country's maternal mortality rate stood at 84 deaths per 100,000 live births in 2000 (United Nations Children Fund, 2015). Under-five mortality was 69 deaths per 1,000 live births in 2000.[2] About one in 14 children died before reaching the age of five, one in 10 in rural Upper Egypt (El-Zanaty & Way, 2001).

Egypt's healthcare system additionally manifested significant inaccessibility and inequities in the access to, the use of and the outcomes of health care. In terms of access, less than 40 percent of Egyptians were insured under HIO in 1997. In terms of equity, the geographic, income and gender healthcare disparities widened during the 1990s. The geographic maldistribution of beds and physicians further worsened geographic differences in outpatient visits and hospital admission rates. Disparities in infant and child mortality and maternal mortality were estimated at 3 to 1 and 5 to 1, respectively, across governorates. Only 16 percent of public spending for health was allocated to the lowest income quintile group. Gender inequalities, especially in access to reproductive health care, were also pronounced (World Bank, 2004).

The healthcare system was profoundly inefficient as well, combining financing and provision functions. Moreover, PHC facilities were underutilized and public hospitals had low occupancy rates despite the surplus of physicians (125,000 licensed physicians in 1997) and hospital beds (2.1 per thousand population in 1997)[3]. Over 60 percent of all PHC visits took place in private facilities in spite of the massive capacity, low cost and physical availability of public PHC facilities in

[1] HIO is a financer and provider of care as well.
[2] Data refers to the 10-year period preceding the Egypt DHS 2000.
[3] Egypt has more beds per capita than other comparable low middle-income countries.

Egypt. The underutilization of PHC services was induced by the poor quality of healthcare provided, the unavailability of drugs and the inappropriate de facto opening times of most public PHC facilities that offered a morning shift only. Hospital occupancy rate was less than 50 percent in the 1990s. In parallel, private out-of-pocket expenditures were substantial (World Bank, 2004).

The quality of public PHC facilities was alarmingly poor, especially in rural areas. The majority of facilities across the country were underfunded, inadequately maintained and equipped and lacked supplies and drugs. Health personnel were inadequately trained.

In addition, allocative and technical inefficiencies posed a threat to the long-term financial stability of the health system, particularly in light of expanding insurance coverage and high population growth.

The GOE launched a comprehensive HSRP in 1997 to address the fundamental challenges in the healthcare system. The main objective of the program is to provide coverage of a basic package of healthcare services for all Egyptians based on five guiding principles: universality, quality, equity, efficiency and sustainability. In this regard, the HSRP introduced an integrated package of service delivery and financing interventions to address the means by which PHC is financed, delivered, organized and managed.

The HSRP came into operation in 2000. The simultaneous implementation of a comprehensive program across Egypt was deemed infeasible due to pre-existent constraints in the healthcare system and the complex nature of interventions to be introduced. Hence, the GOE decided to implement the HSRP over phases. The main targeting took place at the district level. The master plans of governorates relied on a social vulnerability index to target districts of the most vulnerable populations. Early entrants to the program included a group of PHC facilities in Alexandria, Menoufia and Sohag governorates, which represent urban governorates, Lower Egypt and Upper Egypt, respectively. The three pilot governorates represent the three major regions in Egypt. Other governorates followed subsequently.

2.2.1 Service Delivery Interventions

The main objective of the service delivery component of the HSRP is to address persistent needs in maternal and child health through a family health approach to the provision of PHC. A family health model (FHM) for patient care was first adopted in Egypt. The new model allows for rapid integration of family health services where PHC services are provided under the same roof for the entire family. The service delivery component of the HSRP introduced three main supply-side interventions: renewal of PHC infrastructure and equipment, development of human resource and quality assurance.

The HSRP requires each participating healthcare facility to possess a standard catalogue of physical equipment that enables it to provide the basic benefit package (BBP) of healthcare services introduced under the reform. However, a facility has to be accredited first to receive investments in infrastructure and equipment. Human resource development is centered on family practice training for physicians and nurses and subject-specific training for other non-medical specialists in the facilities. To raise the efficiency of often overstaffed facilities and improve the quality of health personnel, the HSRP sets guidelines regarding staff numbers and qualifications required at each public PHC facility. The guidelines also regulate the number of families that could be registered per physician (a maximum of 700 families). The guidelines further regulate the working hours of facilities to ensure that 24-hour services are provided in both rural and urban PHC facilities.

To ensure high quality of care, the HSRP introduced a facility accreditation system and, later, a performance-based financing (PBF) scheme of healthcare providers.[4] MOH established the Quality Improvement and Accreditation Program (QIP) in 1998 to develop both a standardized process of accreditation as well as a system to ensure the continuous improvement and the effective monitoring of facilities' performance. The Quality and Regulation Unit (QRU) in MOH assesses public PHC facilities based on a standardized survey and observation tools given by the QIP. The assessment covers eight dimensions: patient rights, patient care, safety, management of support services, management of information, quality improvement program, family practice and management of the facility. For each facility, the QRU

[4] PBF is also known as pay-for-performance (P4P).

calculates a percent score of compliance to reflect adherence to pre-established accreditation standards. Facilities that achieve a score between 50 and 80 percent or above 80 percent are accredited for one year or two years, respectively.

Accreditation qualifies health facilities to benefit from infrastructure, equipment and human resource interventions under the HSRP. The QRU regularly inspects accredited facilities. Contracted facilities that participate in the financing component of the HSRP are further subject to the autonomous supervision of the relevant regional Family Health Fund (FHF). FHFs supervise the administrative and financial arrangements of facilities. On the central level, the central FHF monitors and evaluates the adherence of regional FHFs to the HSRP strategies at the regional level.

Unreformed facilities are only subject to supervision by the Health District and the Health Directorate. All public PHC facilities –accredited or not– are subject to direct administrative supervision at the governorate level by the Health District and the Health Directorate. Public PHC facilities are routinely visited, inspected and followed up on compliance with standards by supervisors from both entities.

The facilities that are subject to service delivery interventions under the HSRP and are successful in receiving accreditation are referred to as *"accredited"* facilities. A proportion of the accredited facilities became subject to the financing interventions introduced under the HSRP.

2.2.2 Financing Interventions

The financing component of the HSRP includes two main interventions: rechanneling of funds from direct to performance-based financing (supply-side intervention) and instituting a non-linear price system for the uninsured at the point of delivery (demand-side intervention). In parallel, MOH developed a cost-effective BBP to maximize the health benefits per each Egyptian pound (EGP) spent.

In 1999, FHFs were established at the governorate level to separate the financing from the provision of PHC services and to ensure the sustainability of finance of PHC services. FHFs act as intermediate contractors that purchase PHC services from healthcare providers[5] on behalf of uninsured and insured beneficiaries. The FHF

[5] Public, private and NGO.

establishes and supervises the rules and eligibility criteria for PHC providers, contracts with providers provided accreditation and adherence to the FHF's guidelines, rechannels the funds allocated to PHC services by receiving finance from a mix of sources and pays salaries of staff.

A PBF scheme was later integrated in the HSRP in 2001. According to this scheme, FHFs paid monthly incentives to contracted healthcare providers, who deliver the BBP, based on pre-defined performance criteria. To qualify for financial incentives, facilities were required to underscore pre-determined standards of 11 indicators that covered: number of visits per day per physician, number of drugs per visit, rate of patient referral to the district hospital, rate of completion of visit encounter forms, patient satisfaction rate, rate of completion of medical records data, years of protection provided by contraceptive methods, number of children fully vaccinated in the catchment area, patient waiting time, number of ANC visits per pregnant woman and adherence to medical protocols. The indicators were selected to reflect various aspects of service provision. A weight and a required standard were assigned to each indicator.

The higher the performance of a facility with regard to the 11 indicators, the higher the bonuses that were paid by the FHF to this facility. The health personnel in awarded facilities received incentives of 50-275 percent of their base salary, based on job type, years of experience, academic qualifications and on-the-job performance. However, if a reformed facility performed averagely, the PBF incentives it received became more stringent than the non-PBF incentives an unreformed facility received, making the latter facilities more attractive for staff. To ensure the financial sustainability of FHFs, PBF was replaced by the conventional "fee-for-service" mechanism by the end of 2008.

FHFs are financed through different sources, mainly revenues from the new price system (roster fees, visit fees and copayments) and contributions of MOH on behalf of the uninsured. Other main sources include HIO's reimbursements on behalf of its insured individuals who choose to use reformed PHC facilities as well as official donations from internal and external agencies. FHFs use these financial resources to pay monthly incentives to contracted providers based on pre-defined performance

criteria, salaries for contracted personnel and other administrative and operating costs.

The new non-linear price system requires the uninsured to pay a one-off copayment of EGP10 to register in the contracted facility and open a family folder, a copayment of EGP5 (EGP10 in Menoufia) for annual renewal of the family health folder, a copayment of EGP3 per examination and 35 percent of the medical treatment (drugs and other therapy). The poor are officially exempt from any copayments at the point of service in contracted facilities. Some population categories are also exempt such as under-18 orphans without a supporter, divorced women, widows and the unemployed. However, there are concerns over the functioning of exemptions. The main concern is that the majority of individuals have never heard of the payment exemption of the poor (World Bank, 2010).

The new copayment regime became effective in contracted PHC facilities in all governorates and in all MOH PHC facilities in Menoufia. The insured in HIO could also use healthcare services provided by contracted health facilities that are later reimbursed for treatment. Unreformed PHC facilities charge just EGP1 per examination and nothing for registration, treatment/therapy or even drugs. However, the *de facto* examination fees charged by many facilities are above the official fee scale, especially in the rural areas.

The PHC facilities that are subject to the financing interventions under the HSRP are referred to as *"contracted"* facilities as they enter into contractual agreements with the respective FHF. However, facilities do not necessarily transfer from the first phase (accreditation) to the second phase (contracting) of the HSRP.

Figure 2.1 illustrates the pathway of Egypt's HSRP. Out of 4,882 eligible PHC facilities across Egypt, a total of 2,549 facilities were successfully accredited by year 2014, 763 out of which became contracted by 2014. A total of 1,786 PHC facilities remained accredited only by 2014. These facilities have not been subject to any of the financing interventions introduced under the HSRP (Figure 2.2).

Figure 2.1: The pathway of the HSRP

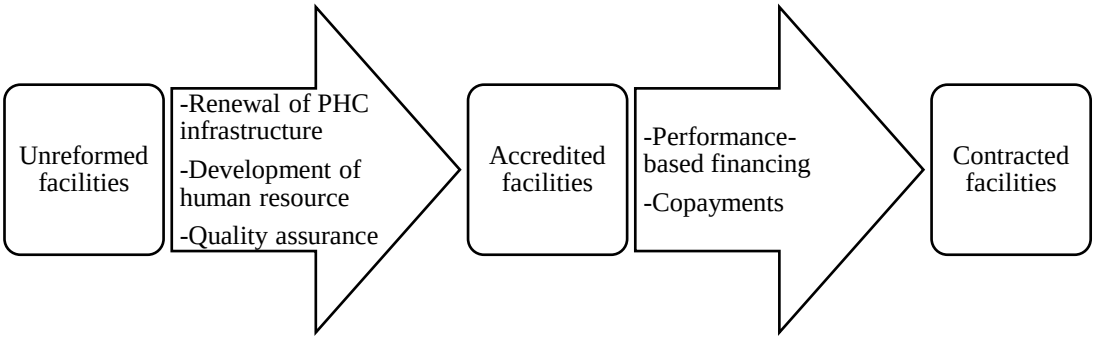
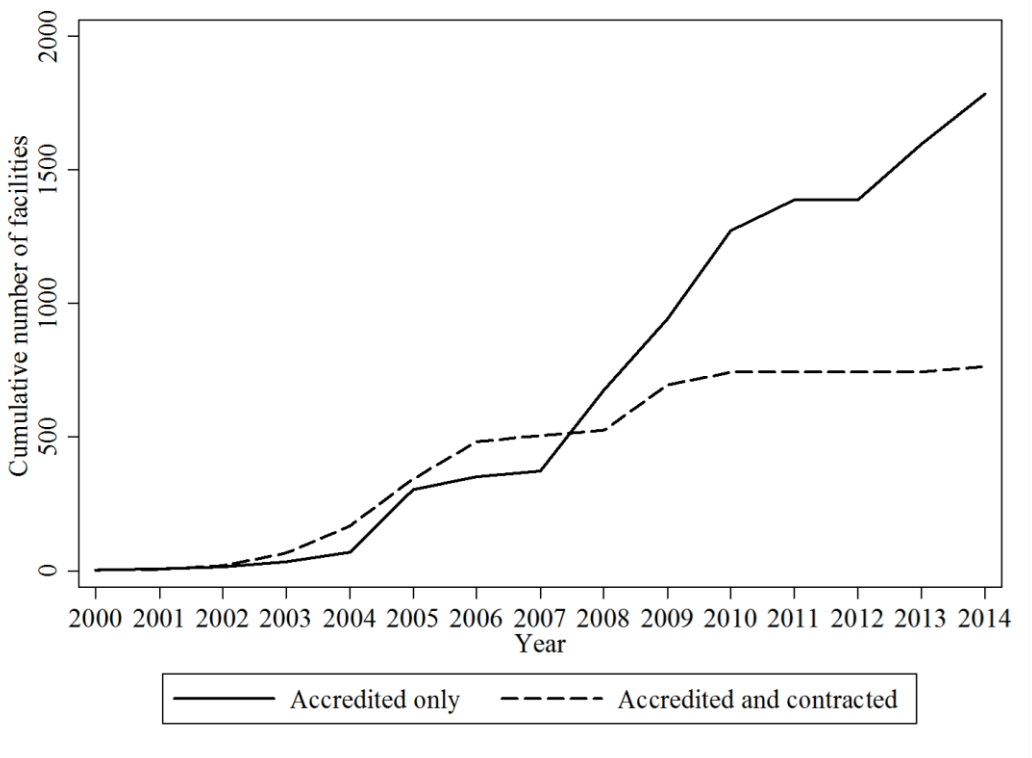


Figure 2.2: Progress of the HSRP's implementation



Source: Author's calculations based on data from Egypt's MOH

A Quasi-Natural Experiment. A quasi-natural experiment is associated with the implementation of Egypt's HSRP. As discussed earlier, the HSRP targeting is not random but rather relies on a social vulnerability index to target districts of the most vulnerable populations. However, we highlight the fact that the main targeting takes place at the district level rather than village or facility level. Decentralized and

autonomous allocation of funds takes place within districts. The Health District and the Health Directorate are responsible for such allocation within each district. This allocation could be random or follows criteria we do not observe. We also highlight the fact that, even at the district level, the HSRP targeting does not strictly follow the social vulnerability index. Later in the empirical chapters, we would investigate whether the differences in the district characteristics between reformed and unreformed health facilities are consistent with the targeting criteria of the HSRP.

2.3 DATA SOURCES

In this section, we describe the facility-level data, the district-level data, the DHS and the other supplemental sources of information that this study draws on.

2.3.1 Facility-Level Data

Our main explanatory variables reflect the HSRP interventions at the facility level. These variables draw from facility-level data that is collected from Egypt's MOH and associated agencies. The Department of PHC in MOH keeps records of all PHC facilities that are affiliated to the Ministry. The QRU regularly accredits and inspects health facilities. The Family Health Unit (FHU) tracks the contracting of facilities with relevant FHFs. We obtain a list of PHC facilities in all governorates across the country during the period 2000-2014 and merge this data with data on interventions for accreditation and contracting. Table 2.1 disaggregates Egypt's health facilities by the phases of reform. Out of 4,882 facilities eligible to join the HSRP, a total of 2,549 (52 percent) and 763 (16 percent) facilities are accredited and contracted, respectively, by year 2014 (Table 2.1).

Table 2.1: Health facilities by status of reform

Cumulative number of facilities	Years			
	2000	2005	2008	2014
Total	3,527	3,976	4,839	5,300
Eligible to join the HSRP	3,065	3,533	4,410	4,882
Accredited	3	647	1,200	2,549
Contracted	2	343	525	763

Source: Author's calculations based on data from Egypt's MOH

We also obtain information on the geographic coordinates of all health facilities across Egypt from the National Information Center for Health and Population (NICHP) affiliated to MOH. We use the data on the Global Positioning System (GPS) locations of PHC facilities and the GPS locations of women in our facility-level analyses to link each woman to her nearest mapped facility.

Besides information on the interventions, we obtain detailed information on the characteristics of health facilities from MOH. We use this data to generate a set of facility-level controls. Whenever possible, these controls are included in our facility-level analyses to account for possible differences between treated and control facilities prior to participation in different interventions under the HSRP. The data obtained addresses the labor force, the building condition and the population coverage of facilities.

2.3.2 The DHS

The main source of data used to construct our dependent variables is the Egypt cross-section DHS, conducted in years 2000, 2005, 2008 and 2014. The Egypt DHS provides detailed information on fertility, family planning, infant and child mortality, maternal and child health and nutrition. The Egypt DHS consists of two main questionnaires: the household (HH) questionnaire and the ever-married women (EMW) questionnaire. Table 2.2 summarizes the sample selection, the survey coverage and the response rate of each of the four waves of the Egypt DHS.

The EMW questionnaire, from which we extract our data, provides information on a nationally representative sample of EMW (between the ages of 15-49) during each month of around a five-year period beginning five years prior to the date of the interview. The observations are the EMW and their children. We present the characteristics of the EMW sample for the Egypt DHS waves in Table 2.3.

The four waves of the Egypt DHS allow us to match early and late entrants to the HSRP. Combining data from the four waves of the DHS provides information on family planning and maternal and child health for 73,336 women and 51,219 births during the period 1995-2014.

Table 2.2: Description of the Egypt DHS

	The 2000 Egypt DHS	The 2005 Egypt DHS	The 2008 Egypt DHS	The 2014 Egypt DHS
Sample selection				
(1) Stage 1	500 primary sampling units selected (228 shiakhas/towns and 272 villages)	682 primary sampling units selected (298 shiakhas/towns and 384 villages)	610 primary sampling units selected (275 shiakhas/towns and 335 villages)	884 primary sampling units selected (481 shiakhas/towns and 445 villages before dropping North and South Sinai from the sample)
(2) Stage 2	1,000 segments from the parts in each shiakha/town and village chosen	1,359 segments from the parts in each shiakha/town and village chosen	1,267 segments from the parts in each shiakha/town and village chosen	1,838 segments from the parts in each shiakha/town and village chosen
(3) Stage 3	Random sample of HHs drawn	Random sample of HHs drawn	Random sample of HHs drawn	Random sample of HHs drawn
Survey coverage	17,521 HHs 15,649 women	22,807 HHs 19,565 women	19,739 HHs 16,571 women	29,471 HHs 21,903 women
Response rate	99.5%	99.5%	99.7%	99.4%

Source: The Egypt DHS, 2000, 2005, 2008 and 2014

Table 2.3: Characteristics of the EMW sample of the Egypt DHS

Characteristic of the EMW sample		2000	2005	2008	2014
Demographic and socio-economic characteristics (weighted %)					
Marital status	Married	92.4	93.4	93.2	94.0
	Widowed	5.0	3.9	4.1	3.1
	Divorced/separated	2.7	2.7	2.8	2.9
Age	15-19	4.0	4.1	3.8	3.5
	20-24	14.4	15.2	15.6	14.0
	25-29	18.3	19.4	20.4	21.8
	30-34	17.3	16.5	16.1	19.0
	35-39	17.2	16.4	15.6	16.1
	40-44	14.0	14.7	15.0	13.2
	45-49	14.8	13.7	13.5	12.4
Urban-rural residence	Urban	44.1	41.3	41.2	35.0
	Rural	55.9	58.7	58.8	65.0
Place of residence	Urban governorates	19.2	16.9	17.7	12.7
	Lower Egypt	43.8	43.2	46.1	49.0
	Urban	12.5	11.3	11.7	10.7
	Rural	31.3	31.9	34.4	38.3
	Upper Egypt	35.6	38.8	34.8	37.4
	Urban	11.6	12.4	10.8	11.1
	Rural	24.0	26.4	24.0	26.2
	Frontier governorates	1.3	1.1	1.4	0.9
Education	No education	43.2	34.6	32.1	24.0
	Some primary	13.2	11.3	8.4	6.1
	Primary complete/some secondary	13.0	14.0	14.6	17.4
	Secondary complete/higher	30.5	40.1	44.9	52.4
Educational attainment	Median years of schooling	3.1	6.3	7.6	10.1
Literacy	Percentage literate	50.2	59.1	64.7	73.2
Wealth quintile	Lowest	NA	18.3	18.4	17.9
	Second	NA	19.4	19.7	19.7
	Middle	NA	20.2	20.5	22.2
	Fourth	NA	21.2	21.2	20.9
	Highest	NA	20.9	20.2	19.4
Employment status (%)					
Employment status	Employed in the 12 months preceding the survey				
	Currently employed	16.8	21.5	16.4	15.5
	Not currently employed	0.9	0.7	0.3	0.6
	Not employed in the 12 months preceding the survey	82.3	77.8	83.4	83.9
Occupation	Professional/technical/managerial	40.9	36.0	45.9	45.3
	Clerical	23.6	12.7	11.4	9.2
	Sales and services	11.5	17.3	18.6	21.4
	Skilled manual	5.5	5.0	4.0	6.8
	Unskilled manual	NA	3.3	5.1	1.2
	Agriculture	17.3	25.5	14.5	16.1
Type of employment	Agricultural work	17.3	25.5	14.5	NA
	Non-agricultural work	81.5	74.3	85.1	NA
Characteristics of the HH population					
Dependency ratio		69.2	62.1	61.5	NA
Mean size		5.2	4.9	4.6	4.1
HH headship (%)	Male	88.0	87.7	86.6	87.1
	Female	12.0	12.3	13.4	12.9
Housing characteristics					
Electricity (%)	Yes	97.7	99.4	99.6	99.8
	No	2.3	0.6	0.4	0.2
Source of drinking	Piped into residence/plot	80.8	89.7	91.4	90.9

Characteristic of the EMW sample		2000	2005	2008	2014
water (%)					
Sanitation facility (%)	Modern flush toilet	33.4	43.0	42.2	56.1
	Traditional with tank flush	3.6	1.9	2.3	NA
	Traditional with bucket flush	57.5	52.9	54.7	NA
	Pit toilet/latrine	3.0	1.0	0.3	0.0
	No facility	2.6	1.1	0.4	0.1
Mean rooms per HH		3.7	3.9	3.8	NA
Mean persons per room		1.6	1.8	1.7	NA

NA: Not available.
Source: The Egypt DHS, 2000, 2005, 2008 and 2014

However, it is important to note that the Egypt DHS does not allow us to track the same women and children over time. Therefore, we collapse data from each DHS wave at the facility level to construct facility-level indicators for each wave. These indicators constitute the dependent variables of this study. Thus, we investigate the effects of interventions under the HSRP at the facility level rather than the individual level.

In most recent DHSs, the groupings of HHs that participate in the survey, known as clusters, are geo referenced. These survey cluster coordinates are collected in the field using GPS receivers, usually during the survey sample listing process. In general, the GPS readings for most clusters are accurate to approximately less than 15-20 meters. We obtain the GPS data points of all interviewed women in the Egypt DHS and use this information in the subsequent empirical chapters to spatially link women to their nearest facilities, and then, construct facility-level health outcomes.

2.3.3 The Population and Housing Census

We use data from Egypt's 2006 Population and Housing Census to generate a set of district-level social and economic indicators. Whenever possible, these indicators are included in our analyses to capture possible differences between treated and control facilities prior to the launch of the HSRP. The socio-economic indicators included are the illiteracy ratio, the unemployment ratio, the income dependency ratio, inaccessibility to electricity, inaccessibility to potable water, the average family size, the HH crowding factor and the population size. These eight indicators are used to construct the vulnerability index that is used for the HSRP targeting.

Egypt's Population and Housing Census is the primary source of information on the universe of Egyptian HHs. The census is the most comprehensive source of

demographic, economic and social statistics for the entire population. The Central Agency for Public Mobilization and Statistics (CAPMAS) of Egypt conducts this Population and Housing Census every 10 years.

2.4 CONCLUSION

In this chapter, we provide the contextual setting of our analyses in the four empirical chapters to follow. First, we discuss the fundamental challenges faced by Egypt's healthcare system with respect to health outcomes, access to care, equity, efficiency, quality and financial stability. These are the challenges that triggered the launch of the HSRP in 1997. Under the HSRP, different types of interventions were introduced to improve the service delivery and financing of healthcare services. The service delivery component of the HSRP introduced three supply-side interventions: renewal of PHC infrastructure and equipment, development of human resource and quality assurance. The financing component introduced two main interventions: PBF of healthcare providers on the supply side and user fees on the demand side. Second, we present the data sources of this study. We describe the facility-level data, the district-level data, the Egypt DHS and the other sources of information used to construct our explanatory and dependent variables in the subsequent empirical chapters. Moreover, we summarize the background characteristics of the EMW interviewed whose data is later used to calculate our health outcomes of interest.

3. THE EFFECT OF IMPROVING THE QUALITY OF HEALTH CARE THROUGH ACCREDITATION ON FAMILY PLANNING AND MATERNAL AND CHILD HEALTH

3.1 INTRODUCTION

In 1997, a comprehensive Health Sector Reform Program (HSRP) was launched in Egypt to address fundamental challenges in the healthcare system. As discussed in chapter 2, the HSRP introduced an integrated package of service delivery and financing interventions to address the means by which primary health care (PHC) is financed, delivered, organized and managed. One of the cornerstones of the HSRP is the facility accreditation program. The program is defined as a process for evaluating PHC facilities according to a set of standards that define activities and structures that directly contribute to improved patient outcomes. The main aim of the program is to provide the HSRP with a framework for continuous quality improvement. The results of accreditation are also used as the basis for performance-based contracting with the Family Health Fund (FHF).

Accreditation of healthcare providers has been established in many high-income countries, and some low- and middle-income countries, as an approach to improve the quality of care that combines the two elements of quality assurance and quality improvement (Hort *et al.*, 2013). There is consistent evidence that shows that accreditation programs improve the process of care provided by healthcare services (Alkhenizan & Shaw, 2011). However, there is little evidence on the effectiveness of accreditation in terms of patient-level outcomes in all settings (Hinchcliff *et al.*, 2012; Flodgren *et al.*, 2016; Brubakk *et al.*, 2015). To date, no study in a low- or middle-income country investigated the effect of quality improvement through accreditation on key patient outcomes such as family planning, maternal health and child health outcomes. Since accreditation usually entails a significant cost, investigating its effectiveness is crucial, especially in settings where resources are constrained.

In this chapter, we attempt to fill the gap in the literature by investigating the effect of quality improvement through accreditation on patient outcomes in a middle-income

country. The chapter exploits the quasi-natural experiment associated with the introduction of Egypt's facility accreditation program to estimate the effect of having access to an accredited facility on a set of family planning, maternal health and child health outcomes. To do this, difference-in-differences (DD) is combined with propensity score matching (PSM). Diversified datasets on the health of populations in catchment, the characteristics of health facilities and the socio-economic indicators of districts are used.

The chapter proceeds as follows. Section 3.2 provides background information for the analysis of this study; section 3.3 discusses the econometric methods used in the study; section 3.4 constructs our dependent and explanatory variables; section 3.5 presents the descriptive statistics and the estimation results; section 3.6 discusses several tests of the underlying identification strategy used; and section 3.7 concludes.

3.2 BACKGROUND

In this section, we give an overview of the facility accreditation program, discuss the anticipated effect of the program and review evidence on the effect of quality improvement through accreditation interventions in low- and middle-income countries.

3.2.1 The Facility Accreditation Program

In 1997, the GOE launched the HSRP, a new PHC strategy to reform the health system in phases over a period of 15-20 years. One of the key marketing points of the newly introduced PHC model of service provision is that it would improve access to quality care. To this end, the facility accreditation program was introduced with the aim of improving the quality of PHC services. The program is defined by Egypt's MOH as an organized process to monitor the quality of services and influence the behavior and functions of healthcare providers to ensure compliance with quality standards. The program also allows facilities to improve the quality of services by providing follow-up visits where technical assistance is provided to develop an improvement plan.

Any PHC facility in Egypt is eligible to participate in the facility accreditation program. As part of the HSRP, accreditation is obligatory to all facilities interested in joining the program and contracting with the Family Health Fund. In addition, accreditation is voluntary to any other facility interested in being accredited. For a

facility to become eligible, it must meet specific criteria. The facility must have a process to monitor, evaluate, and improve the quality of care to its patients. In parallel, the facility must have a patient record system. In addition, the facility must provide a defined package of services including reproductive health obstetrics and gynecology, neonatal care, pediatric and adult medical care, basic emergency care and preventive health services. The facility must also provide services that include ambulatory care with or without inpatient services. Finally, the facility must be in operation for at least six months, have appropriate license by MOH and relevant medical union and operates in compliance with all government laws and regulations.

The survey is a key step in the accreditation program. The survey process consists of a site visit to the PHC facility. This visit is conducted by a team of experts trained in accreditation using pre-set accreditation survey instruments and tools. The purpose of the accreditation survey is to evaluate the extent to which a facility complies with the nationally established accreditation standards, and accordingly, determine whether a facility is awarded or denied accreditation.

Eight categories were selected to be included in the assessment: patient rights, patient care, safety, management of support services, management of information, quality improvement program, family practice and management of the facility. Optimal standards in each category were developed focusing on key processes, activities, or outcomes that facilities should achieve.

During the accreditation survey, trained surveyors use three approaches to collect data and measure compliance with the established standards. The three approaches are: record review of specific administrative and clinical records, observation of the performance of specified tasks in specified areas in addition to personal interviews. The data is later analyzed using a computerized accreditation program. If a facility scores from 80 percent to 100 percent of the total survey scores, it is granted full accreditation for a period of two years. If a facility scores between 50 percent and 79 percent of the total survey score, it is granted provisional accreditation for one year, after which a reassessment survey of the facility is conducted to assess improvements made in the problem areas identified in the first accreditation report. If a facility scores less than 50 percent, accreditation is denied.

The scoring criteria that measures the accreditation standards ranges from zero to three. Scores of zero, one, two and three denote that an accreditation standard is not met, unacceptable (partially met), acceptable (partially met) and fully met, respectively. All the scores from each activity are added to get the aggregate for the accreditation standard. The average score for each standard is calculated by dividing the aggregate scores by the frequency of activities. The scores are then weighed at the sub-area score level (level one) and the overall facility score level (level two) as shown in Tables 3.1 and 3.2, respectively.

The quality dimension “patient care” contributes the most to the total accreditation score. This dimension measures the extent to which patients receive appropriate care. This focuses on compliance with clinical practice guidelines as well as appropriate diagnosis, assessment, treatment, follow-up and patient counseling.

The focus of the assessment of the sub-area “antenatal care (ANC)” of the quality dimension “patient care” is to get a general understanding of the quality of ANC at the facility. For example, the surveyor assesses if a comprehensive history and physical examination is performed for all patients. The general physical examination should include weight measurement, height measurement, blood pressure measurement as well as measurement of edema of lower limbs. The surveyor also assesses if the necessary diagnostic tests (laboratory and radiology) are performed on time to determine the diagnosis. These tests include but are not limited to blood analysis, complete urine analysis and ultrasound according to clinical guidelines. In addition, the surveyor assesses that all treatment plans are appropriate according to clinical guidelines. For example, supplementation of iron and folic acid in first trimester is checked. The surveyor as well assesses the number of ANC visits according to clinical guidelines. For these visits, there is another assessment if a comprehensive history and physical examination is performed for all patients. In addition, the surveyor assesses if some educational messages are discussed with patient. For example, the physician should assist pregnant women have better knowledge and understanding of their immunization status (tetanus toxoid), the importance and the number of visits that should be made prior to delivery, alarming signs such as bleeding and the delivery services in the facility.

Table 3.1: Sub-area score

Quality dimension	Sub-area	Sub-area weight
Patient rights	Patient rights	2
	Dimension total	**2**
Patient care	General clinical areas	3
	Hypertension	3
	Diabetes	3
	ANC*	3
	Normal delivery, neonatal	3
	Postnatal care	3
	IMCI**	3
	Immunization	3
	Family planning	3
	Dimension total	**27**
Safety	Infection control	3
	Sterilization	3
	Employee health safety	1
	Environmental safety	2
	Dimension total	**9**
Support services	Emergency	2
	Laboratory	2
	Radiology	2
	Pharmacy	3
	Housekeeping	1
	Kitchen	1
	Laundry	1
	Dimension total	**12**
Management of information	Medical records	2
	MIS***/reporting	1
	Dimension total	**3**
Quality improvement program	Quality improvement program	2
	Dimension total	**2**
Family practice model	Prevention and screening	3
	Continuity of care	3
	Referral	3
	Dimension total	**9**
Management of the facility	Human resource development	1
	Management	1
	Budgeting	1
	Continuous education	1
	Provider satisfaction	1
	Dimension total	**5**

*ANC: Antenatal care. **IMCI: Integrated Management of Child Illnesses. ***MIS: Management Information System.
Source: Egypt's MOH

Table 3.2: Overall facility score

Quality dimension	Dimension weight	% of total score
Patient rights	1	6%
Patient care	5	29%
Safety	3	18%
Support services	2	12%
Management of information	1	6%
Quality improvement program	1	6%
Family practice model	3	18%
Management of the facility	1	6%
Total	**17**	**100%**

Source: Egypt's MOH

The focus of the sub-area "Integrated Management of Child Illnesses (IMCI)" is the wellbeing of children under five years of age. The surveyor first assesses if a comprehensive history and physical examination is performed for all sick children according to age of child. For example, a child is checked for cough, diarrhea, sore throat, ear infection and fever. Second, this surveyor assesses if the health providers explain to mothers the classification and treatment and any follow-up steps using clear and simple language. Third, the surveyor assesses if cases requiring diagnostic tests are appropriately referred according to IMCI guidelines, when needed. Finally, the surveyor assesses if the facility provides appropriate prevention and treatment to all sick children according to IMCI guidelines.

The assessment of the sub-area "family planning" mainly focuses on counseling sessions. The surveyor assesses if a comprehensive history and physical examination is performed for all new women. The surveyor also assesses if insertion and removal of method is appropriately performed according to guidelines. In addition, the surveyor assists if the facility has a good information/education/communication (IEC) system. For example, the surveyor checks if the health provider discusses the family planning methods and follow-up steps with the client. The surveyor checks if the client is informed about different methods, mode of action, side effect, how to use a method, cost of method, etc. Finally, the surveyor assesses if the client decided on the appropriate method(s).

Equipment and staff in accredited facilities generally follow a plan that complies with international standards. If needed, accreditation is accompanied by a series of interventions to meet equipment quality standards and strengthen staff's competence

in addressing family health needs. Existing delivery units are transformed to meet the criteria for either Family Health Units (FHU) that provide basic outpatient services or Family Health Centers (FHC) that provide limited specialist outpatient and inpatient services. In most cases, this transformation implies upgrading, renewing or adding modern equipment. Examples of equipment in FHUs include sterilization ovens, delivery chairs and dentist chairs. Examples in FHCs include ultrasounds, x-rays, and hematological and cytological labs. To strengthen staff's competence, equipment interventions are accompanied by a comprehensive training package for facility staff. For physicians and nurses, the package focuses on family health practice. For other non-medical specialists in facilities, such as pharmacists, lab technicians and social workers, the package focuses on subject-specific training. In addition, training is a means to introduce substantial administrative changes in facilities, such as reaching out to and rostering families, and keeping medical records electronically and in family folders.

3.2.2 Anticipated Effect of the Facility Accreditation Program

One important characteristic of healthcare markets is the presence of asymmetric information, where healthcare providers have more information than patients with respect to diagnoses, potential treatments, outcomes of services provided and the payments for these services (Arrow, 1963). When health providers, for their own self-interests, act as "imperfect" agents on behalf of patients, over and under provision of care, as well as variations in quality, become a health concern. Different interventions, including accreditation of providers, have evolved in response to these problems. By subjecting healthcare providers to a formal process that makes them meet pre-determined standards, accreditation is expected to minimize variations in medical practice, eliminate medically inappropriate care, control costs and address the possibility that quality is underprovided (Viswanathan & Salmon, 2000; Akerlof, 1970).

In the context of this study, accreditation is expected to have a primary or direct effect on some maternal health, child health and family planning outcomes, and a secondary or indirect effect on other outcomes. Accreditation of health facilities reflects better compliance with standards defining activities and structures that directly contribute to improved patient outcomes. Thus, accreditation standards established to measure

compliance of facilities in the subareas ANC, IMCI and family planning of the quality dimension "patient care" are expected to have a primary effect on ANC coverage (number of ANC visits), quality of ANC (weight measurement, blood pressure measurement, urine sample collection, blood sample collection and iron supplementation), child morbidity prevalence (acute respiratory infection (ARI), fever and diarrhea) and informed choice of contraceptive methods (knowledge of side effects of contraceptive method used and knowledge of other methods of contraception that could be used). These outcomes reflect some of the standards assessed during the accreditation survey. We expect improvements in these outcomes in accredited compared to non-accredited facilities.

In parallel, accreditation is expected to have a secondary effect on the utilization of antenatal and delivery care services. Quality improvement in accredited facilities introduces an incentive for individuals to increase their demand for healthcare services and seek care at accredited facilities. The effect of this incentive is expected to be more significant with respect to the sub-areas of care that are included in the assessment of the accreditation survey. Thus, we expect having access to an accredited facility to be associated with higher ANC coverage (at least four visits), higher institutional delivery and higher skilled assistance during delivery. This expectation holds given that accredited facilities were not functioning at full capacity prior to accreditation and can increase supply in the short term.

3.2.3 Evidence on the Effect of Accreditation

We searched Academic Search Complete, Business Source Complete, CINAHL, EconLit, E-Journals, Health Policy Reference Center, MEDLINE, PsycINFO and SocINDEX via EBSCO database on May 4, 2017. A combination of the keywords "accreditation" and "health" was used to search the abstracts of studies published in EBSCO's databases. We initially limited the results by excluding non-English studies, studies published before 2000 and studies conducted in high-income settings. This search yielded 104 studies after removing exact duplicates from the results. We screened these studies based on title and abstract, and searched the reference lists of the relevant ones. A total of seven studies were finally selected to be reviewed (see Table 3.3).

Table 3.3: Evidence on the effectiveness of accreditation of healthcare providers

Study	Intervention	Outcome measure	Reported effect on outcome
Al Tehewy *et al.* (2009)	-Accreditation of non-governmental organizations' health units	-Patient satisfaction -Provider satisfaction -Compliance with standards	Positive None Positive
Bukonda *et al.* (2002)	-A national hospital accreditation program	-Compliance with standards	Positive
Cleveland *et al.* (2011)	-Accreditation of basic package of health services	-Availability of information -Shared understanding of priorities and oversight -PBF	Positive Positive Positive
El-Jardali *et al.* (2008)	-A national hospital accreditation program	-Perceived quality of care	Positive
Hong *et al.* (2011)	-Accrediting public clinics a gold star against a checklist of 101 quality indicators -Communication campaign	-Contraceptive supply -Counselling services -Examination services -Management -Compliance with standards	Positive Positive Positive Positive Positive
Quimbo *et al.* (2008)	-Accreditation of public/private healthcare providers -Insurance claims review -Financial incentives to healthcare providers	-Average "vignette" score*	Positive
Salmon *et al.* (2003)	-A province public hospital accreditation program	-Compliance with standards -Nurse perceptions of quality -Patient satisfaction -Patient medication education -Medical record retrieval and accuracy -Medical record completeness -Completeness of peri-operative notes -Labeling of ward stock medications -Hospital sanitation	Positive Positive None None None None None None None

*Vignettes are written case scenarios designed to measure the quality of clinical care by measuring a doctor's ability to properly diagnose and treat patients.

The majority of studies reviewed report on the effect of accreditation on compliance with quality standards (Al Tehewy *et al.*, 2009; Bukonda *et al.*, 2002; Hong *et al.*, 2011; Salmon *et al.*, 2003). All studies suggest that accreditation has a positive effect on compliance with standards in the majority of the assessment's areas. A quasi-experimental study in Egypt found that accredited non-governmental health units had better compliance with quality standards compared with non-accredited units (Al Tehewy *et al.*, 2009). Another study in Egypt on the impact of accrediting public clinics a gold star against a checklist of quality indicators found that providers in Gold Star facilities were more likely to adhere to higher quality practices in counselling and examination than in non-Gold Star facilities (Hong *et al.*, 2011). According to a descriptive study in Zambia, a national hospital accreditation program was associated with significant improvement in compliance of accredited hospitals with standards in overall scores and in seven out of 13 functional areas (Bukonda *et al.*, 2002). In South Africa, Salmon *et al.* (2003) used a randomized control trial to investigate the effect of an accreditation program on public hospitals' processes and outcomes. The study found that the average compliance of accredited hospitals to standards improved significantly, while no significant increase was observed in non-accredited hospitals.

Besides compliance with standards, the majority of the studies reviewed report on the effect of accreditation on quality of care measures. These are, for the most part, not patient health outcomes, but downstream process indicators (Al Tehewy *et al.*, 2009; El-Jardali *et al.*, 2008; Hong *et al.*, 2011; Quimbo *et al.*, 2008; Salmon *et al.*, 2003). Unlike compliance with standards, there is no conclusive evidence on the effect of accreditation on quality of care. While El-Jardali *et al.* (2008), Hong *et al.* (2011) and Quimbo *et al.* (2008) report positive effect of accreditation on different indicators of quality of care, the studies employing more robust study designs report mixed effects. These are Salmon *et al.* (2003) and Al Tehewy *et al.* (2009), which used a randomized controlled trial and a quasi-experimental design, respectively. In a study based on data from hospitals in South Africa, Salmon *et al.* (2003) found little or no effect of a randomized accreditation program on quality measures apart from increases in perception of quality among nurses. In Egypt, Al Tehewy *et al.* (2009) found a positive effect of accreditation of non-governmental health units on patient satisfaction with respect to all areas of health service (cleanliness, waiting area, waiting time and staff performance). As for provider satisfaction, the study found a

positive effect on the overall satisfaction score, but no significant difference in the mean satisfaction score between the accredited and non-accredited units with respect to the social environment, administrative environment and family health model.

In conclusion, the available evidence on the effectiveness of quality improvement through accreditation of healthcare providers is limited and of questionable quality. The evidence investigated in this review suggests that accreditation could improve the process of care provided by different providers. However, evidence is limited on the effectiveness of accreditation on patient outcomes. The review did not identify any study that estimated the effect of accreditation on key patient outcomes, such as family planning, maternal health and child health outcomes. In order to ensure that accreditation brings effective quality improvement practices, there is a need to assess quality based on patient outcomes.

3.3 ECONOMETRIC STRATEGY

In this section, we discuss the two methods used to estimate the effect of Egypt's facility accreditation program: DD and DD PSM. The Global Positioning System (GPS) coordinates for PHC facilities nationwide and populations in catchment allow us to link each woman to the nearest mapped PHC facility. Health outcomes are calculated at the facility level based on information about women in catchment areas of facilities. This data allows the comparison of health outcomes for accredited and non-accredited facilities in 2000 and 2005 and again in 2005 and 2008.

3.3.1 DD

The main methods of impact evaluation are regression (least-squares), PSM, DD, regression discontinuity, instrumental variables (IV) and randomization. These methods except regression address the problem of selection bias. Since the publication of the seminal article Ashenfelter & Card (1985), DD has been applied widely (e.g., Card, 1992; Card & Krueger, 1994; Blundell *et al.*, 1998; Angrist & Krueger, 1999; Finkelstein, 2002).

In this study, we use DD to estimate the effect of accreditation under Egypt's HSRP. The DD setup is appropriate for this study as we observe both baseline and post-intervention health outcomes for both accredited and non-accredited health facilities. We compare the health outcomes of accredited facilities (treatment group) and non-

accredited facilities (control group) before and after accreditation. Using this method, we seek to remove biases in post-treatment comparisons between accredited and non-accredited facilities that could result from permanent differences between facilities. We also seek to remove biases that could arise from comparisons between baseline and follow-up years. These are the biases that result from trends.

For each health facility i at time t, we estimate the following DD specification:

$$y_{it} = \alpha + \beta\, acc_{it} + \gamma\, d_{post} + \delta\, acc_{it} * d_{post} + \zeta\, fac_i + \eta\, dist_i + \varepsilon_{it} \qquad (3.1)$$

where y_{it} denotes a health outcome of interest y for facility i at time t. t = 0 for the baseline year (2000 or 2005) and t = 1 for the follow-up year (2005 or 2008). acc_{it} is a treatment dummy variable for an accredited facility i. acc_{it} = 1 if facility i is accredited and acc_{it} = 0 if facility i is non-accredited. The coefficient β captures baseline differences between accredited and non-accredited facilities prior to accreditation. d_{post} is a time-period dummy variable for the follow-up year. d_{post} = 0 for the baseline year and d_{post} = 1 for the follow-up year. The coefficient γ captures any time trends in health outcomes and aggregates factors that could have induced changes in outcomes of interest y even if the facility accreditation program was not introduced. The interaction term $(acc_{it} * d_{post})$ measures treatment after the baseline year. The term is equal to zero for all facilities in the baseline year and for non-accredited facilities in the follow-up year. Only for accredited facilities in the follow-up year, the term is equal to one. δ is the coefficient of interest. It captures the effect of accreditation on each respective outcome at the facility level. fac_i is a vector of facility-level controls that reflect different characteristics of facility i, such as the labor force, the building condition and the population coverage of facilities. To ensure that the differential effect between accredited and non-accredited facilities are attributable to accreditation, we must control for observable facility characteristics that could explain part of these effect. The coefficient ζ captures the effect of these characteristics on health outcomes at the facility level. $dist_i$ is a vector of district-level controls that reflect the social, economic and demographic characteristics of the district in which facility i is located. We include these indicators to reflect the

selection criteria of the targeting of the HSRP and ensure proper matching of treatment and control groups. Thus, the coefficient η captures the effect of district-level social, economic and demographic indicators on health outcomes at the facility level. The error term ε_{it} denotes the unobserved component of facility i at time t. The term summarizes variations in health outcomes not captured by the remainder of covariates.

Our parameter of interest δ is calculated within a regression framework, where for each facility i at time t, $\hat{\delta}$ is given by:

$$\hat{\delta} = \left(y_{ipost}^{T} - y_{ipre}^{T}\right) - \left(y_{jpost}^{C} - y_{jpre}^{C}\right) \tag{3.2}$$

where the term $\left(y_{ipost}^{T} - y_{ipre}^{T}\right)$ denotes the change in health outcomes of treated facilities after being subject to accreditation. The term $\left(y_{jpost}^{C} - y_{jpre}^{C}\right)$ denotes the change in health outcomes of control facilities after being subject to accreditation.

The identifying assumption of $\hat{\delta}$ is that, conditional on the measured covariates, the differences in health outcomes between accredited and non-accredited facilities before and after accreditation would have been the same in the absence of treatment. To test the validity of this assumption, we run several robustness checks later in this chapter (see section 3.6).

For each health outcome, we report the results of estimating three specifications of equation (3.1) for two study periods: 2000-2005 and 2005-2008. Each specification progressively adds controls to test the stability of the coefficients. The DD model's specification (1) does not include any controls. Facility-level controls are added in the model's specification (2). In specification (3), we control for all observable facility- and district-level characteristics that could be correlated with our health outcomes. The bootstrap method is used to estimate the standard errors in all regressions.

We have two econometric concerns. First, differences in the observable characteristics between accredited and non-accredited facilities, which are correlated with our outcomes of interest, could explain the differences in health outcomes between facilities after the reform. The econometric specification we propose accounts for this possibility by controlling for the characteristics that could be correlated with our

health outcomes. We include a set of facility-level characteristics that provide detailed information on the labor force, the building condition and the population coverage of facilities. These are the characteristics that are potentially correlated with our outcomes of interest.

Our second econometric concern is that the regional targeting of the HSRP across districts is not random. As discussed in section 2.2, the targeting follows a social vulnerability index. To eliminate any potential bias, we control for the observed differences at the district level that could explain differential effect. We use a set of district-level social and economic indicators that reflect the selection criteria of the targeting of the HSRP. These indicators are included as control variables in our econometric analyses to ensure proper matching of treatment and control groups. We realize that additional targeting could have taken place within districts, at the facility level. However, within-district targeting lies within the discretion of the district health management and does not follow a criterion. The best option to control for this targeting is to include a set of facility-level characteristics as controls in our analyses whenever applicable.

3.3.2 DD PSM

Endogeneity is the fundamental concern that arises when we estimate the effect of participating in the facility accreditation program on health outcomes. There are some potential sources of endogeneity in our context. Self-selection, for instance, is one potential source of endogeneity. Accredited and non-accredited facilities are likely to differ along observable characteristics, such as labor force, and along other characteristics that are usually unobserved, such as managerial competence of facilities. As noted earlier, the targeting of the reform interventions under the HSRP is not random, but rather follows an observable socio-economic vulnerability index. This non-random assignment to treatment could bias our estimates if the outcomes of the two groups of facilities are compared without accounting for selectivity.

Another potential source of endogeneity is omitted-variable bias. This bias could arise when a facility in a district is self-selected into the HSRP based on unobservable characteristics. Targeting could also depend on non-observables. If this is the case, unobserved facility characteristics could simultaneously affect our dependent variables (i.e., health outcomes) and main explanatory variable (i.e., participation in

the facility accreditation program). If so, the treatment variable is said to be "endogenous" and its estimated coefficient is expected to be biased.

To address these endogeneity issues, we combine DD with the PSM approach. In general, matching is used to identify a comparison group of non-participants with similar pre-intervention characteristics as the treatment group. PSM is a way of determining which characteristics should be used and what weight should be assigned to each characteristic. Instead of matching on a multitude of dimensions in a vector of observable characteristics *Z*, Rosenbaum & Rubin (1983) argues that it is only necessary to match on a single dimension P(*Z*), which is the propensity score. A propensity score is given by:

$$P(Z) = \text{Probability}(\text{Treatment}=1 \mid Z) \tag{3.3}$$

where *Z* is a vector of pre-treatment characteristics that could as well include the pre-treatment value of the health outcome. Treated units are matched to comparison or control units with similar values of P(*Z*).

In this study, we extend the conventional DD estimate by defining health outcomes conditional on propensity scores and applying semi-parametric methods to construct the differences. First, we match treated and control health facilities based on pre-treatment observable characteristics. The intuition behind this matching is to construct a comparison group of non-accredited facilities with similar pre-intervention characteristics as that of accredited facilities. There are many ways to undertake the matching. The simplest methods are the one-to-one matching and the nearest n-neighbor matching. A non-parametric method that we use in our analyses is Kernel matching (Caliendo & Kopeinig, 2008). We use Kernel functions to assign weight to the *j*th control facility matched to the *i*th treated facility using PSM as follows:

$$\omega(i,j) = \frac{K(\frac{P_j - P_i}{a_n})}{\sum_{K \in C} K(\frac{P_k - P_i}{a_n})}, \tag{3.4}$$

where a is the bandwidth parameter; K is the Kernel function and P_i and P_j are the propensity scores for treated and control facilities.

In our context, the propensity score is the probability of being targeted by interventions under the HSRP given a set of social and economic indicators. These are the indicators that construct the socio-economic vulnerability index used by the GOE for the HSRP targeting. After matching treated and control facilities, we use a two-sample t-test to check if there are still significant differences in the means of observable characteristics for both groups (Rosenbaum & Rubin, 1983).

Second, we conduct DD estimations in which health outcomes are defined conditional on the propensity score generated earlier. The DD Kernel PSM estimate for each treated facility *i* is calculated as follows:

$$\hat{\delta}_i = \left(y^T_{ipost} - y^T_{ipre}\right) - \sum_{j \in C} \omega(i,j)\left(y^C_{jpost} - y^C_{jpre}\right) \quad (3.5)$$

DD PSM was first proposed by Heckman *et al.* (1998). The intuition behind combining both methods is to preserve the strengths and minimize the weaknesses of each method. To identify treatment effect in a *"selection on observables"* context, PSM could be used to control for observed characteristics. To identify treatment effect in a *"selection on non-observables"* context, DD could be used to control for unobserved characteristics. Thus, the conventional PSM could not be used to account for non-observables that could explain why the GOE chooses to enroll a facility in a district in the facility accreditation program and that could also affect health outcomes. However, combining DD with matching could at least account for unobserved differences that are constant over time. In addition, it is worth noting that the DD PSM estimates are superior to the conventional DD estimates as no functional form restrictions are imposed when estimating the conditional expectation of the outcome variable using DD PSM.

While we use district-level social and economic indicators to estimate the propensity score, facility-level characteristics are used as additional covariates later in the DD estimations. For each of our health outcomes, we report the results for two study periods: 2000-2005 and 2005-2008.

3.4 DATA

In this section, we present the sources of our data and construct the dependent and explanatory variables included in our analyses. As discussed in chapter 2, we rely on three waves of the Egypt Demographic and Health Survey (DHS) to calculate facility-level health outcomes. We draw on facility-level data from MOH to reflect accreditation under the HSRP and to match accredited and non-accredited facilities. Data from Egypt's 2006 Population and Housing Census conducted by the Central Agency for Public Mobilization and Statistics (CAPMAS) is used to construct a set of district-level social and economic controls.

3.4.1 Dependent Variables

We make use of all the relevant data made available by the Egypt DHS on family planning and maternal and child health. This is the data that we expect to reflect the effect of compliance with quality standards, policies and procedures, which are the focus of accreditation assessment. Our dependent variables are outcomes of informed choice of contraceptive methods, ANC, delivery care and child morbidity prevalence. We collapse the Egypt DHS individual responses at the facility level and calculate facility-level outcomes. We, then, combine outcomes of the 2000, 2005 and 2008 waves of the DHS in a panel. These outcomes constitute the dependent variables of our analyses.

Data Spatial Join. For each wave of the Egypt DHS, we use the GPS coordinates of both women interviewed in the Egypt DHS and health facilities to link each woman to the nearest mapped facility. The idea is to identify women who live in the catchment area of accredited facilities (treatment group) and women who live in the catchment area of non-accredited facilities (control group). The distance matrix tool in Quantum GIS 2.8.2 is used to do the spatial join. We calculate the linear distance between each woman in a DHS cluster and all PHC facilities in Egypt. A woman (i.e., DHS cluster) is then linked to the facility with the smallest distance. So, for each woman, we have an attribute that identifies the nearest facility and the linear distance to this facility. All eligible PHC facilities across Egypt are used during the joining process.

A limitation of this spatial join is that an error could arise if some women seek health care from alternative sources apart from their nearest PHC facility. Examples of these sources are private hospitals and clinics. However, the movement between public and

private facilities is unlikely given that prices are lower in public facilities. Some health services are even provided free-of-charge in public facilities. An error could also arise if despite additional time and financial costs, some women bypass their closest PHC facilities in favor of higher quality PHC facilities or even public hospitals (Akin & Hutchinson, 1999). Again, this bypassing is unlikely in the context of this study as for a woman to use a MOH public PHC facility, she is obliged by MOH to use only the facility in catchment.

Variables Construction. We use Stata 12.0 to recode and calculate health outcomes at the facility level for each of the Egypt DHS waves and combine outcomes in a panel. Our outcomes cover informed choice of contraceptive methods, ANC, delivery care and child morbidity prevalence.

Family Planning. As part of the "family planning" sub-area of the patient care dimension of quality assessed by the accreditation survey, the surveyor checks if the facility has a good IEC system. In this regard, the surveyor checks if the health provider explains to client about the family planning methods and follow-up steps using clear and simple language. Two of the topics that should be discussed with the client are observed in the DHS: side effect of the contraceptive method used and other methods of contraception that could be used.

Thus, we include two family planning outcomes that capture the effect of accreditation on informed choice of contraceptive methods. We calculate the percentage of current users of selected contraceptive methods who were informed of the side effect or problems of the method used *(contsid)*. Women who receive information on the efficacy and side effect of contraceptives used tend to have higher continuation rates than those who do not receive such information. We also calculate the percentage of current users of selected contraceptive methods who were informed of other methods of contraception that could be used *(contoth)*. Informed choice emphasizes that women choose the method that best satisfies their personal and reproductive health needs based on a thorough understanding of other methods of contraception they could use.

ANC. As part of the "ANC" sub-area of the patient care dimension of quality assessed by the accreditation survey, the surveyor checks if physical examination is performed for all patients. Two of these examinations are observed in the DHS: weight

measurement and blood pressure measurement. The surveyor also checks if the necessary diagnostic tests are performed on time to determine the diagnosis. Two of these tests are observed in the DHS: urine sample collection and blood sample collection. In parallel, the surveyor checks if all treatment plans are appropriate according to clinical guidelines. In this regard, the surveyor checks if iron is being supplemented in first trimester.

Thus, we include five ANC outcomes that capture the effect of accreditation on the quality of ANC. We calculate the percentages of mothers who received the following components of ANC: weight measurement *(ancwght)*, blood pressure measurement *(ancbp)*, urine sample collection *(ancus)* and blood sample collection *(ancbs)*. We also calculate the percentage of mothers who received iron supplements *(anciron)* during pregnancy. These components are also highlighted by the World Health Organization (WHO) guidelines that are specific on the content of ANC visits. An outcome of an ANC component, say *ancwght*, is the quotient of the number of women who received ANC for their last birth and who had their weight measured, divided by the number of women with a birth in the last five years who received ANC for their last birth, expressed as a percentage.

As part of the accreditation survey, the surveyor checks as well if the number of ANC visits is according to clinical guidelines. The Egypt DHS data allows us to calculate an outcome of ANC coverage (at least four visits) *(anc4)*. WHO defines *anc4* as "the percentage of women aged 15-49 with a live birth in a given time period who received ANC four or more times." *anc4* is used as a global preferred outcome of access to and use of health care during pregnancy to track performance in maternal health programs. It is a MDG outcome indicator that tackles progress towards Target 5.B of MDG 5. WHO recommends that a woman receives at least four antenatal visits during a normal pregnancy to ensure that antenatal complications are detected and controlled at the earliest stage. A pregnant woman is expected to receive health interventions during antenatal visits that could be vital to her health and the health of her infant as well.

According to the Guide to DHS Statistics, *anc4* is the quotient of the numbers of women who received ANC for their last birth, according to grouped number of visits (four visits), divided by the number of women with a birth in the last five years, expressed as a percentage. As part of the Egypt DHS EMW questionnaire, each female

respondent is asked how many times she received ANC during each of the pregnancies of her children born in the last five years.

Delivery Care. As part of the "ANC" sub-area of the patient care dimension of quality assessed by the accreditation survey, the surveyor assesses patient's knowledge and understanding of delivery services provided in the facility. In an interview, the patient is asked if anyone explained to her these services. The Egypt DHS allows us to calculate two delivery care outcomes to capture the effect of compliance with the accreditation standards in this regard: institutional delivery *(delplac)* and skilled assistance during delivery *(delassist)*. The two outcomes are widely advocated for reducing maternal, perinatal and neonatal mortality. *delplac* captures the effect of accreditation on expanding access to childbirth facilities (particularly public health facilities). *delplac* is also a proxy measure of maternal and neonatal morbidity and mortality. Women who give birth at a health facility are more likely to receive proper medical attention and care during delivery. Their infants as well are more likely to receive proper care after delivery. According to the Guide to DHS Statistics, *delplac* is the quotient of the numbers of live births whose deliveries took place in a health facility, divided by the number of live births in the last five years, expressed as a percentage.

The second but most important measure of delivery care included in our analyses is *delassist.* WHO defines *delassist* as "the proportion of births attended by skilled health personnel." It is a MDG outcome indicator that tackles progress towards Target 5.A of MDG 5, which is to "reduce by three quarters, between 1990 and 2015, the maternal mortality ratio." Empirical literature provides evidence that wider access to professional care during pregnancy and childbirth reduces maternal mortality. Women assisted by skilled health personnel during delivery are less likely to die from any cause related to or aggravated by childbirth (Graham *et al.*, 2001). According to the Guide to DHS Statistics, *delassist* is the quotient of the number of live births assisted by medical provider (doctor or nurse/midwife) during delivery divided by the number of live births in the last five years, expressed as a percent. As part of the Egypt DHS EMW questionnaire, each female respondent is asked either a health professional (doctor or nurse/midwife) or other person (daya or other) or no one assisted with the delivery of each of her children.

Child Morbidity Prevalence. As part of the "IMCI" sub-area of the patient care dimension of quality assessed by the accreditation survey, the surveyor checks if child is checked for cough, diarrhea, sore throat, ear infection and fever. The prevalence of three out of these five illnesses could be calculated using the Egypt DHS data: ARI *(childari),* fever *(childfev)* and diarrhea *(childdiarr).*

ARI is the leading infectious cause of death in children worldwide. Mortality due to ARI accounted for 16 percent of the total deaths among under-five children in 2015. Diarrheal diseases are the second leading cause of death in children under age five (World Health Organization, 2016). The risk of under-five mortality could be diminished substantially through reducing the prevalence of ARI and diarrheal diseases and encouraging women to seek treatment for their children at a health facility or from a healthcare provider. We are not able, however, to calculate indicators of child morbidity treatment as observations in our sample are not statistically sufficient.

We calculate child morbidity prevalence as the quotient of the number of children ill with a cough accompanied by short and rapid breathing (for *childari*), or ill with a fever (for *childfev*) or ill with diarrhea (for *childdiarr*) at any time during the two weeks preceding the interview, divided by the number of children under five years of age, expressed as a percentage. As part of the Egypt DHS EMW questionnaire, each female respondent is asked if any of her children became ill with a fever at any time in the last two weeks, if he/she had an illness with a cough at any time in the last two weeks, if he/she breathed faster than usual with short and rapid breaths or had difficulty breathing when he/she had an illness with a cough and if any of her children had diarrhea in the last two weeks.

3.4.2 Explanatory Variables

The explanatory variables included in the analyses of this study are a treatment variable that reflects participation in the facility accreditation program, facility-level controls, district-level social and economic controls and regional dummies to control for regional discrepancies.

Treatment. The gradual uptake of the facility accreditation program by health facilities provides a quasi-natural experiment with treated and control facilities. A facility is treated if it is accredited and non-treated otherwise. Health facilities that are

subject to additional interventions under the HSRP such as performance-based financing (PBF) and copayments are removed from the dataset.

Controls. Whenever possible, we include a set of facility and district characteristics in our analyses to eliminate potential unobserved heterogeneity, that is, account for possible differences between accredited and non-accredited facilities prior to accreditation. At the facility level, we obtain information on the labor force, the building condition and the population coverage. With respect to labor force, we include the numbers of eight types of workers in a health facility: practitioners (*pract*), specialists (*spec*), pharmacists (*pharm*), nurses (*nurs*), lab technicians (*labtech*), x-ray technicians (*xraytech*), health observers (*obs*) and social workers (*socwork*). With respect to infrastructure, a dummy variable is included to describe the condition of a facility building (*inf*) as bad, average or good. With respect to population, we report the size of population in the catchment areas of facilities (*cov*).

At the district level, we include eight social and economic indicators: the illiteracy ratio (*illit*), the unemployment ratio (*unemp*), the income dependency ratio (*incdep*), inaccessibility to electricity (*elect*), inaccessibility to potable water (*wat*), the average family size (*famsiz*), the household (HH) crowding factor (*crowd*) and the population size (*pop*). In addition, regional dummies are defined for fully urban governorates (*urb*), urban Lower Egypt (*lowurb*), rural Lower Egypt (*lowrur*), urban Upper Egypt (*uppurb*), rural Upper Egypt (*upprur*) and frontier governorates (*front*). These district-level covariates control for both the selection criteria of the HSRP targeting and the demographic variation across districts. As discussed in chapter 2, the regional targeting of the HSRP follows a socio-economic vulnerability index that is constructed from the eight social and economic indicators outlined earlier.

3.4.3 Summary of Variables

The dependent variables included in our analyses are facility-level family planning, maternal health and child health outcomes. Our explanatory variables include a treatment indicator, facility-level controls, district-level controls and regional dummies. Table 3.4 summarizes the description and sources of all variables included in the analyses of this study.

Table 3.4: Description and sources of the variables used

Variable	Description	Measure	Unit	Source	Year(s)
Dependent variables*					
Family planning					
contsid	Proportion of current users of selected contraceptive methods informed of side effect of the method used	Family planning	Percent	Author's calculations based on Egypt DHS	2000, 2005, 2008
contoth	Proportion of current users of selected contraceptive methods informed of other methods of contraception that could be used	Family planning	Percent	Author's calculations based on Egypt DHS	2000, 2005, 2008
Maternal health					
anc4	Proportion of women who received four or more ANC visits	ANC	Percent	Author's calculations based on Egypt DHS	2000, 2005, 2008
ancwght	Proportion of women whose weight was measured as an ANC component	ANC	Percent	Author's calculations based on Egypt DHS	2000, 2005, 2008
ancbp	Proportion of women whose blood pressure was measured as an ANC component	ANC	Percent	Author's calculations based on Egypt DHS	2000, 2005, 2008
ancus	Proportion of women whose urine sample was taken as an ANC component	ANC	Percent	Author's calculations based on Egypt DHS	2000, 2005, 2008
ancbs	Proportion of women whose blood sample was taken as an ANC component	ANC	Percent	Author's calculations based on Egypt DHS	2000, 2005, 2008
anciron	Proportion of women who received iron supplements as an ANC component	ANC	Percent	Author's calculations based on Egypt DHS	2000, 2005, 2008
delplac	Proportion of live births delivered in a health facility	Delivery care	Percent	Author's calculations based on Egypt DHS	2000, 2005, 2008
delassist	Proportion of live births whose delivery was assisted by skilled health personnel	Delivery care	Percent	Author's calculations based on Egypt DHS	2000, 2005, 2008
Child health					
childari	Proportion of under-five children with symptoms of ARI	Child morbidity prevalence	Percent	Author's calculations based on Egypt DHS	2000, 2005, 2008
childfev	Proportion of under-five children with fever	Child morbidity prevalence	Percent	Author's calculations based on Egypt DHS	2000, 2005, 2008
childdiarr	Proportion of under-five children with diarrhea	Child morbidity prevalence	Percent	Author's calculations based on Egypt DHS	2000, 2005, 2008
(2) Explanatory variables					
Treatment					
$treat_{it}$	$treat_{it}$= 1 if a facility is accredited, $treat_{it}$= 0 if a facility is non-accredited	Accreditation	Dummy	Author's calculations based on data from Egypt's MOH	2000, 2008

Variable	Description	Measure	Unit	Source	Year(s)
Facility characteristics					
pract	Practitioners	Labor force	Number	MOH	2005
spec	Specialists	Labor force	Number	MOH	2005
pharm	Pharmacists	Labor force	Number	MOH	2005
nurs	Nurses	Labor force	Number	MOH	2005
labtech	Lab technicians	Labor force	Number	MOH	2005
xraytech	X-ray technicians	Labor force	Number	MOH	2005
obs	Health observers	Labor force	Number	MOH	2005
socwork	Social workers	Labor force	Number	MOH	2005
inf	Building condition of a facility defined as bad, average or good	Infrastructure	Dummy	MOH	2005
cov	Size of the population in catchment served by a facility	Coverage	Number	MOH	2005
District socio-economic indicators**					
illit	Percentage of population (aged 10 years and above) who are illiterate	Illiteracy	Ratio	CAPMAS	2006
unemp	Percentage of population (aged 15 years and above) who are unemployed	Unemployment	Ratio	CAPMAS	2006
incdep	Ratio of the dependent population (aged below 15 years and over 65 years) to the working population (aged 15-64 years)	Income dependency	Ratio	Author's calculations based on CAPMAS's census	2006
elect	Share of the population with no access to electricity	Accessibility to electricity	Ratio	Author's calculations based on CAPMAS's census	2006
wat	Share of the population with no access to potable water	Accessibility to potable water	Ratio	Author's calculations based on CAPMAS's census	2006
famsiz	The number of individuals divided by the number of families	Family size	Ratio	CAPMAS	2006
crowd	Average number of individuals per room	HH overcrowding	Ratio	CAPMAS	2006
pop	Size of population resident in a particular district	Population size	Number	CAPMAS	2006
Regional dummies					
urb	Urban	Regional variation	Dummy	CAPMAS	2006
lowurb	Lower urban	Regional variation	Dummy	CAPMAS	2006
lowrur	Lower rural	Regional variation	Dummy	CAPMAS	2006
uppurb	Upper urban	Regional variation	Dummy	CAPMAS	2006
upprur	Upper rural	Regional variation	Dummy	CAPMAS	2006
front	Frontier	Regional variation	Dummy	CAPMAS	2006
Year dummy					
d_{post}	d_{post}= 1 for follow-up year, d_{post}=0 for baseline year	Follow-up year	Dummy	-	2000, 2005, 2008

*Definitions of dependent variables are obtained from WHO. **Definitions of control variables are obtained from CAPMAS.

3.5 RESULTS

In this section, we first present the descriptive statistics of the sample used in our analyses. Second, we discuss the estimated effect of accreditation on family planning and maternal and child health. As noted earlier, we report the results of estimating three specifications for each health outcome. The DD model's specification (1) includes no controls. Only facility-level controls are included in specification (2). The model's specification (3) includes both facility- and district-level controls. In addition, we report the results of a DD PSM model specification denoted by specification (4). This is our preferred specification as it includes all the explanatory variables listed in Table 3.4 either as matchers or as additional covariates. The analysis of estimation effect is mainly based on the results of our model's specification (4).

3.5.1 Descriptive Statistics

Table 3.5 summarizes the descriptive statistics of the facility-level characteristics, the district-level characteristics and the health outcomes of all facilities observed in years 2000, 2005 and 2008. The statistics of facility and district characteristics indicate relatively moderate variability across health facilities and districts for most variables. We do recognize the relatively large variation in some variables, such as the number of specialists (*spec*), the population coverage of facilities (*cov*) and the income dependency (*incdep*) across districts. However, this variation is justifiable given the nationwide nature of our analysis and is tolerable given the relatively moderate sample size we are using.

With respect to health outcomes, Table 3.5 shows that Egypt performs moderately with respect to the family planning indicators of informed choice. The mean proportion of current users of selected contraceptive methods, who were informed of side effects of the method used (*contsid*), is 47 percent. The mean proportion of those who were informed of other methods of contraception that could be used (*contoth*) is 53 percent. Table 3.5 also shows that Egypt has a relatively moderate level of ANC coverage. On average, 54 percent of pregnant women in Egypt report at least four ANC visits (*anc4*). Egypt performs differently with respect to various components of ANC. Women are more likely to be weighted (*ancwght*) and get their blood pressure measured (*ancbp*) during ANC visits, but less likely to receive iron supplements during pregnancy (*anciron*).

Table 3.5: Descriptive statistics

Variable	N	Mean	Standard deviation	Minimum	Maximum
Facility characteristics					
Practitioners	5,007	3.458	4.693	0.000	57.000
Specialists	5,013	0.568	1.675	0.000	22.000
Pharmacists	5,013	4.058	5.589	0.000	44.000
Nurses	5,013	11.683	12.855	0.000	191.000
Lab technicians	5,013	1.317	1.658	0.000	16.000
X-ray technicians	5,013	0.181	0.824	0.000	14.000
Health observers	5,013	1.344	1.500	0.000	17.000
Social workers	5,013	0.433	1.600	0.000	49.000
Building condition	4,410	1.640	0.630	0.000	2.000
Population coverage	4,737	3.245	10.670	0.061	269.310
District characteristics					
Illiteracy	5,661	32.330	10.285	0.000	54.030
Unemployment	5,661	9.646	4.410	0.000	23.550
Income dependency	5,661	5.059	15.322	0.000	87.402
Inaccessibility to electricity	5,661	1.260	3.385	0.000	69.092
Inaccessibility to potable water	5,661	4.369	6.887	0.000	90.204
Family size	5,661	4.330	0.375	2.580	6.170
HH overcrowding	5,661	1.150	0.107	0.840	1.930
Population size	5,661	31.912	16.358	0.005	117.380
Health outcomes					
Family planning					
Knowledge of side effects	1,927	47.271	23.300	0.000	100.000
Knowledge of contraceptives	1,952	52.796	23.312	0.000	100.000
ANC					
4+ visits	2,098	53.936	26.919	0.000	100.000
Weight measurement	1,977	79.020	22.553	0.000	100.000
Blood pressure measurement	1,977	77.326	22.413	0.000	100.000
Urine sample collection	1,977	62.962	25.839	0.000	100.000
Blood sample collection	1,977	64.155	25.242	0.000	100.000
Iron supplementation	2,100	39.318	23.552	0.000	100.000
Delivery care					
Institutional delivery	2,095	61.515	29.349	0.000	100.000
Skilled-assisted delivery	2,098	71.906	26.674	0.000	100.000
Child morbidity prevalence					
ARI	2,354	7.373	9.116	0.000	71.429
Fever	2,354	12.848	11.512	0.000	80.000
Diarrhea	2,354	8.809	9.937	0.000	63.636

N denotes the number of observations.

In parallel, Egypt has a fairly high level of delivery care coverage through access to health facilities and skilled health personnel. On average, over 60 percent of women deliver their most recent birth in an institutional setting (*delplac*). Over 70 percent of births are assisted by skilled health personnel (*delassist*) (Table 3.5).

Tables 3.6 and 3.7 highlight differences in the facility and district characteristics, respectively, between accredited and non-accredited facilities. We use the two-sample t-test to check whether the means of the two groups differ significantly. Table 3.6 indicates that accredited facilities have a significantly lower number of nurses (*nurs*) during the period 2000-2005 compared to non-accredited facilities. We do not observe significant differences during this study period in the other nine facility-level characteristics. However, we observe that practitioners (*pract*), pharmacists (*pharm*), nurses (*nurs*) and social workers (*socwork*) are significantly less in non-accredited facilities compared to accredited facilities during the period 2005-2008 (Table 3.6).

Table 3.6: Two-sample t-test of facility characteristics of accredited and non-accredited facilities

	2000-2005			2005-2008		
	Non-accredited	Accredited	Difference	Non-accredited	Accredited	Difference
Practitioners	3.367	2.600	0.767	3.137	5.539	-2.402***
			(0.836)			(0.396)
Specialists	0.530	0.833	-0.303	0.522	0.738	-0.216
			(0.298)			(0.144)
Pharmacists	3.931	3.567	0.364	3.723	6.050	-2.327***
			(0.994)			(0.470)
Nurses	11.853	6.400	5.453**	11.937	13.922	-1.985*
			(2.385)			(1.182)
Lab technicians	1.290	0.967	0.323	1.336	1.177	0.159
			(0.300)			(0.146)
X-ray technicians	0.177	0.233	-0.056	0.193	0.163	0.030
			(0.153)			(0.077)
Health observers	1.309	1.167	0.142	1.330	1.348	-0.018
			(0.266)			(0.133)
Social workers	0.437	0.533	-0.097	0.413	0.901	-0.488***
			(0.300)			(0.150)
Building condition	1.630	1.680	-0.050	1.605	1.653	-0.048
			(0.128)			(0.062)
Population coverage	3.305	1.628	1.677	3.269	4.013	-0.745
			(2.068)			(1.007)

Standard errors are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

Table 3.7: Two-sample t-test of district characteristics of accredited and non-accredited facilities

	2000-2005			2005-2008		
	Non-accredited	Accredited	Difference	Non-accredited	Accredited	Difference
Illiteracy	32.376	31.453	0.923	32.672	29.325	3.347***
			(1.875)			(0.884)
Unemployment	9.703	8.883	0.820	9.635	9.688	-0.054
			(0.806)			(0.385)
Income dependency	5.228	0.601	4.627*	5.602	3.386	2.216
			(2.798)			(1.356)
Inaccessibility to electricity	1.285	0.750	0.535	1.357	0.819	0.538*
			(0.624)			(0.308)
Inaccessibility to potable water	4.390	4.294	0.095	4.528	2.593	1.935***
			(1.264)			(0.611)
Family size	4.330	4.405	-0.076	4.334	4.193	0.141***
			(0.068)			(0.032)
HH overcrowding	1.148	1.188	-0.040**	1.141	1.153	-0.013
			(0.019)			(0.009)
Population size	31.646	29.079	2.567	31.072	34.383	-3.311**
			(2.913)			(1.349)

Standard errors are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

On average, districts to which accredited facilities belong are significantly better with respect to only one characteristic during the period 2000-2005: income dependency (*incdep*) (Table 3.7). However, districts to which accredited facilities belong are significantly better with respect to four characteristics during the period 2005-2008: illiteracy (*illit*), inaccessibility to electricity (*elect*), inaccessibility to potable water (*wat*) and the family size (*famsiz*). In addition, accredited facilities are located in districts that are more populated (*pop*) than districts in which non-accredited facilities are located (Table 3.7). Thus, Table 3.7 provides evidence that the actual targeting of the HSRP does not strictly follow the socio-vulnerability index.

3.5.2 Estimated Effects of Accreditation

Table 3.8 provides evidence that accreditation does not have a significant effect on family planning during the study period 2000-2005. However, having access to an accredited facility is associated with a higher likelihood of being informed of the side effects of contraceptives *(contsid)* during the period 2005-2008. Table 3.9 shows that the proportion of women with access to accredited facilities, who are informed of the side effects of contraceptives used *(contsid)*, increased significantly by 10 percentage points (ppts) compared to women with access to non-accredited facilities.

Table 3.8: Estimated effects of accreditation, 2000-2005

	Outcome	DD			DD PSM
		(1)	(2)	(3)	(4)
Family planning	Knowledge of side effects	14.480	12.560	13.110	7.293
		(9.480)	(11.530)	(11.280)	(5.097)
	Knowledge of contraceptives	8.818	11.070	10.920	6.320
		(7.132)	(11.380)	(11.830)	(4.719)
ANC	4+ visits	-7.399	3.702	5.465	2.529
		(12.041)	(9.574)	(8.561)	(3.692)
	Weight measurement	-9.870	-6.371	-4.565	-0.081
		(8.998)	(8.863)	(9.253)	(4.051)
	Blood pressure measurement	-9.997	-5.910	-3.247	0.155
		(9.221)	(9.556)	(9.506)	(3.440)
	Urine sample collection	-15.140	-13.320	-9.640	-5.170
		(10.188)	(11.820)	(11.600)	(3.851)
	Blood sample collection	-15.094	-6.236	-2.618	2.831
		(11.976)	(12.960)	(12.070)	(4.354)
	Iron supplementation	11.155	15.530**	16.120**	11.540***
		(7.926)	(6.971)	(7.501)	(3.171)
Delivery care	Institutional delivery	-0.628	10.290	13.740*	8.187**
		(10.771)	(7.320)	(7.238)	(3.890)
	Skilled-assisted delivery	2.734	10.400	12.320	10.510**
		(10.156)	(8.850)	(7.927)	(4.097)
Child morbidity prevalence	ARI	-4.696	-6.769*	-6.744*	-9.211***
		(3.077)	(3.503)	(3.643)	(1.565)
	Fever	-2.907	-5.286	-5.232	-8.571***
		(3.767)	(4.568)	(4.884)	(1.949)
	Diarrhea	0.487	3.922	3.588	-0.302
		(3.917)	(4.893)	(5.030)	(1.581)

Each row represents a separate regression. Standard errors are reported in parentheses. Bootstrapped standard errors are reported for specifications (1) to (3). *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

In addition, we find that accreditation has limited significant effect on ANC during both study periods. The estimates of 4+ visits *(anc4)*, weight measurement *(ancwght)*, blood pressure measurement *(ancbp)*, urine sample collection *(ancus)* and blood sample collection *(ancbs)* are statistically insignificant during the period 2000-2005. However, having access to an accredited facility is associated with a higher likelihood of receiving iron supplements during pregnancy *(anciron)*. Table 3.8 shows that the proportion of women with access to accredited facilities, who receive iron supplements during pregnancy *(anciron)*, increased significantly by 12 ppts between 2000 and 2005 compared to women with access to non-accredited facilities.

Table 3.9: Estimated effects of accreditation, 2005-2008

	Outcome	DD			DD PSM
		(1)	(2)	(3)	(4)
Family planning	Knowledge of side effects	16.583***	12.790**	13.630***	10.330**
		(4.030)	(5.337)	(4.554)	(4.237)
	Knowledge of contraceptives	9.843**	6.015	6.434	-0.703
		(4.189)	(4.862)	(4.512)	(4.035)
ANC	4+ visits	-2.639	-4.188	-3.663	-2.301
		(4.235)	(3.816)	(3.949)	(3.549)
	Weight measurement	4.949**	3.630	4.548	5.328*
		(2.489)	(3.500)	(3.601)	(2.729)
	Blood pressure measurement	4.278	1.288	1.805	3.870
		(2.775)	(3.563)	(3.836)	(2.969)
	Urine sample collection	6.342	3.910	4.772	3.661
		(4.913)	(4.606)	(4.873)	(3.762)
	Blood sample collection	-0.241	-4.318	-3.388	-5.532
		(4.593)	(5.036)	(5.171)	(4.039)
	Iron supplementation	3.009	-0.523	0.611	1.622
		(4.806)	(4.259)	(4.173)	(3.839)
Delivery care	Institutional delivery	-4.495	-3.766	-1.364	-0.272
		(3.998)	(5.636)	(4.427)	(4.398)
	Skilled-assisted delivery	-1.626	-0.546	1.501	3.880
		(4.057)	(4.878)	(3.756)	(4.082)
Child morbidity prevalence	ARI	1.449	2.639	3.087	2.710
		(1.792)	(2.332)	(2.418)	(1.736)
	Fever	2.008	1.342	1.775	0.278
		(1.947)	(2.357)	(2.445)	(2.003)
	Diarrhea	-0.589	0.783	1.121	-1.684
		(1.862)	(2.626)	(2.537)	(1.972)

Each row represents a separate regression. Standard errors are reported in parentheses. Bootstrapped standard errors are reported for specifications (1) to (3). *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

We also find that having access to an accredited facility is associated with a higher likelihood of weight measurement during ANC visits *(ancwght)* during the period 2005-2008. Table 3.9 shows that the proportion of women with access to accredited facilities, who have their weight measured during ANC visits *(ancwght)*, increased significantly by 5 ppts between 2005 and 2008 compared to women with access to non-accredited facilities. However, the estimates of 4+ visits *(anc4)*, blood pressure measurement *(ancbp)*, urine sample collection *(ancus)*, blood sample collection *(ancbs)* and iron supplementation *(anciron)* are statistically insignificant during this period.

More importantly, we observe that accreditation has a significant positive effect on delivery care during the study period 2000-2005. Having access to an accredited facility is associated with a higher likelihood of both institutional delivery *(delplac)* and skilled assistance during delivery *(delassist)* during the period 2000-2005. Table

3.8 indicates that institutional delivery *(delplac)* and skilled assistance during delivery *(delassist)* increased by more than 8 ppts and 11 ppts, respectively, among women with access to accredited facilities. However, the estimates of institutional delivery *(delplac)* and skilled assistance during delivery *(delassist)* are statistically insignificant during the period 2005-2008. In parallel, we observe that accreditation has a significant positive effect on child morbidity prevalence during the study period 2000-2005. Table 3.8 indicates that accreditation reduced the prevalence of both childhood ARI *(childari)* and childhood fever *(childfev)* among children with access to accredited facilities by more than 9 ppts and 8 ppts, respectively, compared to children with access to non-accredited facilities. However, we do not observe any significant effect on all child morbidity prevalence outcomes during the period 2005-2008.

A comparison between the estimates in Tables 3.8 and 3.9 indicates that the positive effect of the facility accreditation program is not intense during the study period 2005-2008 compared to the period 2000-2005. Some of these effects are even reversed. The change in the effect is justified by the fact that interventions under the HSRP have been slowing down and weakening since 2005. This trend becomes more apparent when we check the extent to which facilities comply to reform rather than the rate by which facilities join the HSRP. A plausible indicator of compliance is the accreditation score. While more facilities get accredited, we observe that accreditation scores were increasing until 2004 but started to decrease since then. Accreditation compliance also varies across governorates (Grun & Ayala, 2006). A high level of commitment to and participation in the HSRP was evident in the preparation and early implementation phases. However, successive changes in the leadership of the healthcare sector in Egypt affect the ownership of and commitment to reform efforts.

3.6 ROBUSTNESS CHECKS

In this section, we run several tests to ensure the robustness of our main results of estimated effects obtained in the previous section. First, we run placebo tests to verify the functional form of our DD set-up. Second, the quality of matching facilities is assessed using two-sample t-tests. Finally, we inspect the sensitivity of our DD PSM estimates.

3.6.1 Placebo Test

We run a placebo test by defining a "false" lagged accreditation intervention. If the functional form of the DD set-up is properly specified, pre-accreditation estimations would yield null results. That is, the facility accreditation program should not have any effect on the health outcomes of accredited facilities before being subject to accreditation.

We use the data of the study period 2000-2005 to verify the results of the period 2005-2008. For the period 2000-2005, facilities that are accredited after 2005 are defined as treated and facilities that are not accredited after 2005 are defined as control. Facilities that are subject to additional interventions under the HSRP are removed from the dataset.

The results of our placebo test are reported in Table 3.10. The treatment estimates are not significantly different from zero for all health outcomes. That is, differences between accredited and non-accredited facilities reported in section 3.5 only emerged after the introduction of the facility accreditation program. I.e., accreditation causes the effects observed rather than the other way around.

3.6.2 Quality of Matching

To check the extent to which observable characteristics are balanced in the matched sample, we use the balancing two-sample t-test of the difference in means of covariates across matched samples of facilities. Our covariates of interest are the ones used earlier to match treated and control health facilities. The results of the t-test are reported in Table 3.11. As the table indicates, there are no systematic differences in general at the baseline in the means of observed characteristics between accredited and non-accredited facilities. That is, matching on the propensity score is successful.

Table 3.10: Estimated effects of placebo accreditation, 2000-2005

	Outcome	Difference (baseline)	Difference (follow-up)	DD PSM
Family planning	Knowledge of side effects	-12.705*	-2.819	9.886
				(8.009)
	Knowledge of contraceptives	-6.589	0.060	6.649
				(8.082)
ANC	4+ visits	4.180	-4.955	-9.134
				(7.874)
	Weight measurement	-0.050	2.016	2.066
				(7.214)
	Blood pressure measurement	2.863	0.186	-2.678
				(7.193)
	Urine sample collection	-2.420	1.809	4.229
				(8.668)
	Blood sample collection	4.270	-1.740	-6.011
				(9.683)
	Iron supplementation	3.817	8.212	4.394
				(9.004)
Delivery care	Institutional delivery	-4.584	-10.182	-5.598
				(6.534)
	Skilled-assisted delivery	-6.477	-15.995***	-9.518
				(5.922)
Child morbidity prevalence	ARI	4.972	6.501*	1.530
				(3.644)
	Fever	3.107	7.327	4.220
				(5.522)
	Diarrhea	-0.691	-2.898	-2.207
				(3.987)

Each row represents a separate regression. The covariates are the facility characteristics, district socio-economic indicators and regional dummies from Table 3.4. Standard errors are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

3.6.3 Sensitivity of Results

We further inspect the sensitivity of our results to the type of the Kernel function, the bandwidth of the Kernel function and the estimation method of the propensity score. To do the Kernel matching, we first must specify the type of the Kernel function. We initially use the epanechnikov Kernel to obtain our main results. This is the default type. In Tables 3.12 and 3.13, we compare the main results of the estimated effects reported to the results obtained based on other types of functions (gaussian, biweight, uniform and tricube). In general, we find that our main estimation results are not sensitive to the type of the Kernel function.

Table 3.11: Balancing two-sample t-test, 2005-2008

	Difference in mean covariates between accredited and non-accredited facilities							
	Illiteracy	Un-employment	Income dependency	Inaccessibility to electricity	Inaccessibility to potable water	Family size	HH overcrowding	Population size
Family planning								
Knowledge of side effects	-1.014	-0.038	-0.510	-0.064	-0.343	-0.042	0.003	1.689
Knowledge of contraceptives	-1.014	-0.038	-0.510	-0.064	-0.343	-0.042	0.003	1.689
ANC								
4+ visits	-1.090	-0.029	-0.666	-0.095	-0.405	-0.051*	-0.000	2.082
Weight measurement	-1.242	0.044	-0.705	-0.097	-0.424	-0.048	-0.002	2.334*
Blood pressure measurement	-1.242	0.044	-0.705	-0.097	-0.424	-0.048	-0.002	2.334*
Urine sample collection	-1.242	0.044	-0.705	-0.097	-0.424	-0.048	-0.002	2.334*
Blood sample collection	-1.242	0.044	-0.705	-0.097	-0.424	-0.048	-0.002	2.334*
Iron supplementation	-1.090	-0.029	-0.666	-0.095	-0.405	-0.051*	-0.000	2.082
Delivery care								
Institutional delivery	-1.131	-0.028	-0.676	-0.091	-0.406	-0.052*	-0.000	1.967
Skilled-assisted delivery	-1.098	-0.027	-0.656	-0.094	-0.400	-0.051*	-0.000	2.100
Child morbidity prevalence								
ARI	-1.223	-0.017	-0.597	-0.079	-0.342	-0.051*	0.000	1.822
Fever	-1.223	-0.017	-0.597	-0.079	-0.342	-0.051*	0.000	1.822
Diarrhea	-1.223	-0.017	-0.597	-0.079	-0.342	-0.051*	0.000	1.822

Means and t-test are estimated by linear regression. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

Table 3.12: Sensitivity to the type of the Kernel function, 2000-2005

	Outcome	Main results	Type of function			
			Gaussian	Biweight	Uniform	Tricube
Family planning	Knowledge of side effects	7.293	7.662	6.820	7.973	7.859
		(5.097)	(5.079)	(5.085)	(5.091)	(5.105)
	Knowledge of contraceptives	6.320	6.788	5.704	6.884	6.863
		(4.719)	(4.729)	(4.711)	(4.715)	(4.718)
ANC	4+ visits	2.529	1.483	3.417	1.666	1.732
		(3.692)	(3.764)	(3.648)	(3.737)	(3.722)
	Weight measurement	-0.081	-3.173	0.760	-1.254	-1.008
		(4.051)	(4.250)	(4.018)	(4.103)	(4.090)
	Blood pressure measurement	0.155	-2.687	1.099	-1.911	-1.051
		(3.440)	(3.595)	(3.381)	(3.573)	(3.519)
	Urine sample collection	-5.170	-9.939**	-3.476	-7.927**	-7.082*
		(3.851)	(4.021)	(3.827)	(3.911)	(3.888)
	Blood sample collection	2.831	-1.749	4.794	-0.215	0.718
		(4.354)	(4.554)	(4.316)	(4.440)	(4.407)
	Iron supplementation	11.540***	11.010***	11.304***	11.634***	11.684***
		(3.171)	(3.215)	(3.140)	(3.201)	(3.192)
Delivery care	Institutional delivery	8.187**	7.568*	8.840**	8.085**	7.665*
		(3.890)	(3.957)	(3.839)	(3.958)	(3.929)
	Skilled-assisted delivery	10.510**	9.623**	11.083***	9.984**	10.040**
		(4.097)	(4.179)	(4.023)	(4.183)	(4.152)
Child morbidity prevalence	ARI	-9.211***	-9.232***	-9.186***	-9.282***	-9.242***
		(1.565)	(1.559)	(1.568)	(1.562)	(1.563)
	Fever	-8.571***	-8.526***	-8.613***	-8.546***	-8.529***
		(1.949)	(1.937)	(1.955)	(1.943)	(1.945)
	Diarrhea	-0.302	-0.304	-0.331	-0.250	-0.237
		(1.581)	(1.578)	(1.582)	(1.582)	(1.582)

Each row represents a separate regression. Standard errors are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

To do the Kernel matching, we also must specify the bandwidth of the Kernel function. The choice of bandwidth implies a trade-off between bias and efficiency. On one hand, a small bandwidth decreases the bias of estimates as we use the most similar observations to construct the counterfactual. The characteristics of these facilities are, in general, very similar. However, a small bandwidth decreases the efficiency of estimates as we ignore a lot of information from the sample. The fact that many control facilities are not used for the estimation implies an increase in the imprecision of estimates caused by a higher variance. On the other hand, a large bandwidth increases both the bias and efficiency of estimates. The bandwidth choice is, therefore, a compromise between a small variance and an unbiased estimate of the true density function. This choice is more important in practice than the choice of the type of the Kernel function (e.g., Silverman, 1986; Pagan & Ullah, 1999).

Table 3.13: Sensitivity to the type of the Kernel function, 2005-2008

	Outcome	Main results	Type of function			
			Gaussian	Biweight	Uniform	Tricube
Family planning	Knowledge of side effects	10.330** (4.237)	11.413*** (4.310)	9.808** (4.229)	11.543*** (4.245)	10.879** (4.244)
	Knowledge of contraceptives	-0.703 (4.035)	0.488 (4.152)	-1.206 (4.020)	0.277 (4.057)	-0.132 (4.050)
ANC	4+ visits	-2.301 (3.549)	-1.924 (3.666)	-2.526 (3.543)	-1.981 (3.555)	-1.959 (3.553)
	Weight measurement	5.328* (2.729)	4.899* (2.833)	5.411** (2.717)	5.127* (2.753)	5.211* (2.747)
	Blood pressure measurement	3.870 (2.969)	3.467 (3.060)	4.039 (2.961)	3.675 (2.984)	3.696 (2.982)
	Urine sample collection	3.661 (3.762)	4.359 (3.860)	3.649 (3.755)	3.911 (3.773)	3.779 (3.773)
	Blood sample collection	-5.532 (4.039)	-5.512 (4.157)	-5.715 (4.030)	-5.226 (4.059)	-5.232 (4.054)
	Iron supplementation	1.622 (3.839)	2.219 (3.954)	1.854 (3.832)	1.306 (3.844)	1.473 (3.842)
Delivery care	Institutional delivery	-0.272 (4.398)	0.865 (4.574)	-0.391 (4.382)	-0.196 (4.432)	-0.203 (4.418)
	Skilled-assisted delivery	3.880 (4.082)	4.531 (4.259)	3.930 (4.064)	3.546 (4.116)	3.743 (4.104)
Child morbidity prevalence	ARI	2.710 (1.736)	2.831 (1.789)	2.679 (1.737)	2.732 (1.742)	2.735 (1.736)
	Fever	0.278 (2.003)	0.368 (2.095)	0.248 (2.005)	0.477 (2.012)	0.330 (2.003)
	Diarrhea	-1.684 (1.972)	-0.731 (2.023)	-1.777 (1.972)	-1.424 (1.983)	-1.547 (1.973)

Each row represents a separate regression. Standard errors are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

The default bandwidth of the Kernel function initially used to obtain our main results is 0.06. Alternative bandwidths are tried (bandwidths = 0.05 and 0.1). Tables 3.14 and 3.15 show our main results of the estimated effects using different bandwidths. We find that our main results are not sensitive in general to the bandwidth parameter.

The estimation of propensity score depends on a parametric specification (commonly logit or probit). We specify probit estimation of the propensity score to obtain our main estimation results. This is the default used by Stata 12.0. The specification of the propensity score equation affects the quality of matching and, consequently, the results. Therefore, we specify logit estimation of the propensity score and re-run the DD PSM models to test the sensitivity of our results to the estimation method of the propensity score. The results of this exercise are reported in Table 3.16. We find that the estimates for both methods of estimation match for most outcomes.

Table 3.14: Sensitivity to the bandwidth of the Kernel function, 2000-2005

	Outcome	Main results	Bandwidth	
			0.05	0.1
Family planning	Knowledge of side effects	7.293	6.736	7.631
		(5.097)	(5.083)	(5.083)
	Knowledge of contraceptives	6.320	5.569	6.736
		(4.719)	(4.707)	(4.727)
ANC	4+ visits	2.529	3.480	1.755
		(3.692)	(3.633)	(3.760)
	Weight measurement	-0.081	0.769	-1.497
		(4.051)	(4.022)	(4.096)
	Blood pressure measurement	0.155	1.218	-1.851
		(3.440)	(3.377)	(3.542)
	Urine sample collection	-5.170	-3.456	-8.374**
		(3.851)	(3.827)	(3.901)
	Blood sample collection	2.831	4.732	-0.673
		(4.354)	(4.315)	(4.435)
	Iron supplementation	11.540***	11.684***	10.923***
		(3.171)	(3.132)	(3.214)
Delivery care	Institutional delivery	8.187**	8.960**	7.838**
		(3.890)	(3.815)	(3.959)
	Skilled-assisted delivery	10.510**	11.284***	9.778**
		(4.097)	(4.002)	(4.173)
Child morbidity prevalence	ARI	-9.211***	-9.133***	-9.243***
		(1.565)	(1.566)	(1.559)
	Fever	-8.571***	-8.563***	-8.545***
		(1.949)	(1.953)	(1.936)
	Diarrhea	-0.302	-0.318	-0.320
		(1.581)	(1.582)	(1.578)

Each row represents a separate regression. Standard errors are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

In conclusion, the previous robustness checks rule out an existing trend that could challenge the DD PSM identifying assumptions. Our robustness checks also provide evidence that the main estimation results reported in section 3.5 are not sensitive in general to alternative types of the Kernel function, bandwidths of the Kernel function and estimation methods of the propensity score.

Table 3.15: Sensitivity to the bandwidth of the Kernel function, 2005-2008

	Outcome	Main results	Bandwidth	
			0.05	0.1
Family planning	Knowledge of side effects	10.330**	10.097**	11.158***
		(4.237)	(4.233)	(4.267)
	Knowledge of contraceptives	-0.703	-1.085	0.174
		(4.035)	(4.023)	(4.085)
ANC	4+ visits	-2.301	-2.584	-1.947
		(3.549)	(3.549)	(3.601)
	Weight measurement	5.328*	5.457**	4.816*
		(2.729)	(2.717)	(2.810)
	Blood pressure measurement	3.870	4.060	3.823
		(2.969)	(2.962)	(3.043)
	Urine sample collection	3.661	3.669	4.892
		(3.762)	(3.758)	(3.846)
	Blood sample collection	-5.532	-5.687	-4.769
		(4.039)	(4.033)	(4.132)
	Iron supplementation	1.622	1.784	2.431
		(3.839)	(3.839)	(3.901)
Delivery care	Institutional delivery	-0.272	-0.307	0.273
		(4.398)	(4.383)	(4.484)
	Skilled-assisted delivery	3.880	4.011	4.104
		(4.082)	(4.067)	(4.180)
Child morbidity prevalence	ARI	2.710	2.693	2.699
		(1.736)	(1.737)	(1.764)
	Fever	0.278	0.225	0.138
		(2.003)	(2.005)	(2.069)
	Diarrhea	-1.684	-1.799	-0.890
		(1.972)	(1.973)	(1.999)

Each row represents a separate regression. Standard errors are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

Table 3.16: Sensitivity to the estimation method of the propensity score

	Outcome	2000-2005		2005-2008	
		Probit	Logit	Probit	Logit
Family planning	Knowledge of side effects	7.293	7.524	10.330**	10.428**
		(5.097)	(5.014)	(4.237)	(4.241)
	Knowledge of contraceptives	6.320	5.466	-0.703	-0.720
		(4.719)	(4.649)	(4.035)	(4.022)
ANC	4+ visits	2.529	4.699	-2.301	-2.162
		(3.692)	(3.500)	(3.549)	(3.545)
	Weight measurement	-0.081	0.995	5.328*	5.350
		(4.051)	(3.965)	(2.729)	(2.728)
	Blood pressure measurement	0.155	-0.287	3.870	4.047
		(3.440)	(3.335)	(2.969)	(2.968)
	Urine sample collection	-5.170	-1.385	3.661	3.753
		(3.851)	(3.839)	(3.762)	(3.756)
	Blood sample collection	2.831	7.864*	-5.532	-5.396
		(4.354)	(4.376)	(4.039)	(4.035)
	Iron supplementation	11.540***	6.279**	1.622	1.855
		(3.171)	(2.976)	(3.839)	(3.835)
Delivery care	Institutional delivery	8.187**	7.725**	-0.272	-0.207
		(3.890)	(3.691)	(4.398)	(4.389)
	Skilled-assisted delivery	10.510**	10.715***	3.880	3.808
		(4.097)	(3.815)	(4.082)	(4.079)
Child morbidity prevalence	ARI	-9.211***	-9.123***	2.710	2.710
		(1.565)	(1.562)	(1.736)	(1.735)
	Fever	-8.571***	-8.489***	0.278	0.235
		(1.949)	(1.953)	(2.003)	(2.005)
	Diarrhea	-0.302	-0.168	-1.684	-1.812
		(1.581)	(1.583)	(1.972)	(1.975)

Each row represents a separate regression. Standard errors are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

3.7 CONCLUSION

In this chapter, we use data from the Egypt DHS waves of 2000, 2005 and 2008 to investigate the effect of quality improvement through Egypt's facility accreditation program. We are concerned with the effect on family planning, ANC, delivery care and child morbidity prevalence. DD is combined with Kernel PSM to address the potential for endogeneity bias.

We find evidence that accreditation has multiple positive effects on delivery care and child morbidity prevalence during the study period 2000-2005. Having access to an accredited facility is associated with a higher likelihood of both institutional delivery and skilled assistance during delivery. In parallel, accreditation is associated with lower prevalence of childhood ARI and childhood fever during this period among children with access to accredited facilities. However, we do not observe an effect of accreditation on family planning outcomes during the period 2000-2005. The effect

of accreditation on ANC outcomes as well is limited. The only positive effect captured is with respect to iron supplementation during pregnancy.

We also observe that the positive effects of the facility accreditation program are not intense during the study period 2005-2008 compared to the period 2000-2005. Some of these effects are even reversed. The only positive effects captured during this period are with respect to knowledge of side effects of contraceptive method used and weight measurement during ANC visits. The results of the period 2005-2008 highlight the fact that a high level of commitment, which is a reflection of strong political will, is indispensable for the success of quality improvement interventions in low- or middle-income countries. Decentralization in no way diminishes the necessity of a high level of commitment from the central government.

The findings of this chapter suggest that accreditation as a means for improving the quality of care could be associated with significant improvements in delivery care and child morbidity prevalence in low- and middle-income countries. However, we found that accreditation alone was not sufficient to sustain high quality of care, especially with respect to family planning and ANC. A possible explanation is that the facility accreditation program was successful in improving the process of care provided but did not have the anticipated effect on patient outcomes in accredited facilities. There could also be factors other than accreditation that might have affected performance differentially in accredited and non-accredited facilities. One factor is the nature and effectiveness of outreach activities carried out by facilities. Our results encourage an enquiry in this direction. Moreover, future research on this topic should broaden its scope to investigate which interventions, if combined with accreditation, could be associated with improved patient outcomes. There is evidence that improvements could be achieved, for example, through combining accreditation with properly monitored and well-designed payment or incentive schemes (Quimbo *et al.*, 2008).

4. THE EFFECT OF INTRODUCING USER FEES ON FAMILY PLANNING AND MATERNAL AND CHILD HEALTH

4.1 INTRODUCTION

In 1997, the Government of Egypt (GOE) launched the Health Sector Reform Program (HSRP). As discussed in chapter 2, the main aim of the program is to provide coverage of a basic benefit package (BBP) of health services for all Egyptians. The HSRP has a service delivery component and a financing component. The former focuses on quality improvement through facility accreditation. The latter introduced two interventions, one on the supply-side and one on the demand-side. Funds are re-channeled, on the supply side, from direct to performance-based financing (PBF) of healthcare providers. On the demand side, user fees were introduced in 2003 in public primary healthcare (PHC) facilities participating in the financing component of the HSRP. Uninsured beneficiaries are required to pay registration and renewal fees as well as copayment fees that include visit fees, drug copayment and copayments for other interventions.

The effect of introducing or increasing user fees in low- or middle-income countries is controversial. On one hand, user fees are advocated as an effective means of generating additional revenue and improving the quality of health services. On the other hand, user fees are found to be a financial barrier for the poor who wish to use health services. Several studies investigated the effect of introducing or increasing user fees, however, the effects observed were immediate and abrupt. While it takes several years for health financing interventions to reach full impact, studies did not investigate whether the effects were sustained over the longer term. (Ensor *et al.*, 2017). Moreover, the quality of evidence provided is low (Lagarde & Palmer, 2008).

This chapter contributes to the literature on demand-side financing in health by providing high-quality evidence on the medium-term effect of introducing user fees in a middle-income setting. Using difference-in-differences (DD), we estimate the effect of introducing user fees on the utilization of family planning, antenatal care (ANC) and delivery care services, women's access to health care, and child health

status during the period 2008-2014. Financial incentives to contracted facilities participating in both components of the HSRP were discontinued at the end of 2008. Thus, the main difference between "accredited only" facilities participating in the service delivery component only and contracted facilities is that the latter became authorized to collect user fees from beneficiaries. This allows us to estimate the effect of introducing user fees by comparing the outcomes of "accredited only" facilities and contracted facilities.

The chapter proceeds as follows. Section 4.2 provides background information for the analysis of this study, section 4.3 discusses the econometric method used; section 4.4 constructs our dependent and explanatory variables; section 4.5 reviews the descriptive statistics and presents our main results; section 4.6 reports the results of our robustness checks; and section 4.7 concludes.

4.2 BACKGROUND

In this section, we give an overview of cost sharing under Egypt's HSRP, discuss the anticipated effect of introducing user fees and review evidence on the effect of introducing or increasing user fees in a low or middle-income settings.

4.2.1 Cost Sharing under the HSRP

Ministerial decree 147 of the year 2003 was issued to increase the ability of the Family Health Funds (FHF) to generate revenues by authorizing FHUs and FHCs to collect user fees and drug copayments from beneficiaries. As determined by the decree, uninsured beneficiaries are required to pay 10 Egyptian pounds (EGP) as registration fees for each enrolled person, up to a maximum of EGP30 per family. Uninsured beneficiaries are also required to pay EGP5 for annual renewal of registration for each enrolled person, up to a maximum of EGP15 per family. Besides registration and renewal fees, uninsured beneficiaries are required to pay copayment fees that include visit fees, drug copayment and copayments for other interventions. The decree requires uninsured beneficiaries to pay a copayment of EGP3 per examination and one third of the price of medical treatment (drugs and other therapy). These registration, renewal and copayment fees are applicable in contracted PHC facilities in all governorates except Menoufia. According to Ministerial Decree 231/2006, uninsured beneficiaries in Menoufia are required to pay EGP20 and EGP10 as

registration and renewal fees, respectively, for each enrolled person, without maximum.

The Ministry of Health (MOH) transfers to the FHF 100 percent of the registration fees, the renewal fees and copayments for the uninsured poor beneficiaries, and 25 percent of those fees (including copayments) for the uninsured non-poor beneficiaries. The insured by the Health Insurance Organization (HIO) can also use health services provided by contracted facilities in all governorates. The facilities are subsequently reimbursed for treatment by the HIO.

Although the fee structure covers only a small share of the actual cost of providing a basic benefit package (BBP) of services, enrolled uninsured beneficiaries identified as poor are officially exempt from any user fees at the point of service in contracted facilities in all governorates. Some population categories are also exempt such as under-18 orphans without a supporter, divorced women, widows, the unemployed, etc. However, there are concerns over the functioning of exemptions. The main concern is that the majority of individuals have never heard of the payment exemption of the poor (World Bank, 2010).

As for unreformed PHC facilities, uninsured beneficiaries are charged just EGP1 per examination and nothing for registration or treatment including drugs. However, the de facto examination fees charged by many facilities are above the official fee scale, especially in the rural areas. Beneficiaries who are insured by the HIO pay in unreformed facilities according to the rules established for their coverage. Any fees that are collected go directly to MOH (for the uninsured) and the HIO (for the insured).

4.2.2 Anticipated Effect of Introducing User Fees

Economic theory suggests that introducing user fees are expected to drive demand for health services in two opposite directions. On one hand, the negative price-elasticity of demand suggests that demand is expected to decrease as the price incurred by the consumer increases. On the other hand, the increased willingness to pay for an improved quality of service is expected to increase demand.

Utilization of Health services. The conventional theory of consumer demand suggests that an increase in the price of a good/service is expected to decrease the demand for this good/service. Similarly, introducing user fees is expected to decrease

the demand for health services by increasing the price incurred by the consumer at the time of consumption. The underlying theory of this expected negative effect is Grossman's human capital model of the demand for health. The economic theory of the demand for health and health care primarily stems from the Grossman human capital approach to health (Grossman, 1972; Grossman, 2000). Grossman extended the neoclassical approach to the consumer demand theory to the commodity of health care and drew from the human capital theory [Becker (1964, 1967); Ben-Porath, 1967; Mincer, 1974]. He constructed and estimated a model of demand for the commodity "good health", where the demand for health care is derived from the demand for "good health".

In Grossman's model, the utilization of health care as well as the time consumed while seeking health care both construct the inputs to the household (HH) health production function, according to which, individuals produce gross investment in health. According to Grossman, an increase in the price of a unit of health input (non-time and time) tightens the budget constraint of an individual and make him/her use less health inputs and produce less health. Thus, introducing user fees for health services, especially to the poor, is likely to discourage them from using health services.

Quality of Care. Introducing user fees can have a positive effect on the quality of health services for which thet are charged, which can be associated with higher demand for these services. For this expectation to hold, revenues generated from user fees should be used by the charging facility to finance quality improvements such as maintenance or renewal of the equipment or the facility, or in-service training for health workers. A share of the revenues generated can also be used to incentivize health workers through performance-based financing (PBF) schemes. However, user fees can also lead to overprovision of services, that is, moral hazard on the part of the healthcare provider.

In parallel, introducing user fees can have an indirect positive effect on the quality of services. User fees can provide incentives for users to monitor their providers and to demand better care. As users pay for services, they have an incentive to demand high quality to ensure they get their money's worth. However, it is not clear whether users can judge the quality of the services they receive.

In this context, it is important to note that even if introducing user fees can improve the quality of the services for which fees are charged, this quality improvement is not necessarily translated into higher utilization of these services or better health outcomes of the population. Introducing user fees for health services can cause users to cut their utilization of these services and to divert to alternative providers of low quality or even self-treatment. This change in utilization patterns is expected to have a negative effect on health outcomes.

Even if the quality of services does not actually improve, introducing user fees can attach value to a service, thus increasing demand by increasing perception of quality of services. I.e., user fees may encourage utilization if they are interpreted as a signal of higher quality (Bagwell & Riordan, 1991; Riley, 2001).

Thus, if accompanied by quality improvements, introducing user fees under the HSRP can introduce an incentive for users to increase their demand for health services and seek care at contracted facilities. However, it is unclear if this positive effect will offset the negative effect suggested by the theory of consumer demand.

4.2.3 Evidence on the Effect of Introducing/Increasing User Fees

On May 10, 2017, we used EBSCO to search several databases: Academic Search Complete, Business Source Complete, CINAHL, EconLit, E-Journals, Health Policy Reference Center, MEDLINE, PsycINFO and SocINDEX. We searched the abstracts of studies published in these databases using a combination of the following keywords: “user fee”, “user charge”, “copayment”, “cost sharing”, “cost recovery” or “fee for service”, along with “health”. We initially limited the results by excluding non-English studies, studies published before 2000 and studies conducted in high-income countries. This search yielded 195 studies after removing exact duplicates from the results. The studies were screened based on title and abstract. Only studies on the effect of introducing or increasing user fees were included. Studies on the effect of removing or reducing user fees were excluded. Moreover, studies on the effects of user fees on variables other than patient outcomes or patient-perceived quality of care were excluded. The reference lists of the relevant studies were searched as well. A total of 13 studies were finally selected to be reviewed (see Table 4.1). The majority of studies reviewed report negative effects of introducing or increasing user fees. These negative effects are typically observed with respect to the utilization of services.

Table 4.1: Evidence on the effectiveness of introducing/increasing user fees

Study	Intervention	Outcome measure	Reported effect on outcome
Audibert & Mathonnat (2000)	-Introduction of user fees in all health facilities from basic local units to some national hospitals	-Drug availability -Vaccinations for children less than a year old -Vaccinations for pregnant women -Curative consultations -Contacts per inhabitant	Mixed Positive Positive Positive Positive
Benjamin *et al.* (2001)	-Introduction of user fees for obstetric services in urban clinics and a general hospital -Higher user fees for obstetric services in urban clinics and a general hospital	-Attendance to antenatal clinics on first visits (immediate) -Attendance to antenatal clinics on first visits (12 months) -Institutional delivery	Negative Positive None
Bratt *et al.* (2002)	-Higher user fees in private non-profit clinics	-Gynecology visits -IUD* insertion visits -IUD revisits -ANC visits	Negative Negative Negative Negative
Chawla & Ellis (2000)	-Introduction of direct user charges and indirect insurance payments in government healthcare facilities, accompanied by quality improvements	-Reporting illness -Seeking treatment -Seeking formal treatment	Mixed None Mixed
Cohen & Dupas (2010)	-Higher price at which antenatal clinics sell long-lasting antimalarial insecticide-treated bed nets (ITNs) to pregnant women	-ITN sales -Pregnant women acquiring an ITN -Pregnant women not only acquiring the ITN but also reporting using it at follow-up	None Negative Negative
Issifou & Kremsner (2004)	-Higher consultation fees in a private hospital	-Pediatric outpatient visits -Malaria cases among outpatients	Negative Positive
Jacobs & Price (2004)	-Introduction of user fees at a district referral hospital	-Care seeking from private practitioners -Patients admitted for malaria -Patients admitted for diarrhea/dysentery -Patients admitted for respiratory infections -Patients admitted for dengue	Positive None None None Negative

Study	Intervention	Outcome measure	Reported effect on outcome
		-Patients admitted for deliveries	None
		-Patients admitted for other conditions	None
		-Total patients (non-TB**)	None
		-Mortality rate per admitted patients	Negative
		-Reporting inability to pay costs of hospitalization	Negative
		-Patients admitted	None
		-Pediatric patients admitted	None
	-Higher user fees at a district referral hospital	-Care seeking from private practitioners	Positive
		-Patients admitted for malaria	Negative
		-Patients admitted for diarrhea/dysentery	None
		-Patients admitted for respiratory infections	None
		-Patients admitted for dengue	Negative
		-Patients admitted for deliveries	None
		-Patients admitted for other conditions	Positive
		-Total patients (non-TB)	None
		-Mortality rate per admitted patients	Negative
		-Reporting inability to pay costs of hospitalization	None
		-Patients admitted	None
		-Pediatric patients admitted	None
Kipp *et al.* (2001)	-Introduction of user fees in 38 government health units	-Utilization by outpatients	Negative
		-Utilization by outpatients (urban/semi-urban)	Negative
		-Utilization by outpatients (rural)	Positive
		-Utilization by malaria patients (urban/semi-urban)	Negative
		-Utilization by malaria patients (rural)	Positive
Kremer & Miguel (2007)	-Introduction of user fees for preventive deworming drugs in primary schools	-Utilization of deworming drugs	Negative
Matee & Simon (2000)	-Introduction of user fees for dental health services provided by the government	-Attendance of dental patients	Negative
		-Dental treatment demands	None
		-Dental treatment pattern	None
Mubyazi *et al.* (2006)	-Introduction of user fees in public health facilities	-Quality of care	None
		-Malaria patient attendances	None

Study	Intervention	Outcome measure	Reported effect on outcome
Richard *et al.* (2007)	-Introduction of user fees for emergency obstetric care in an urban district hospital	-Emergency referrals from health centers	Positive
		-Major obstetric interventions (MOI)	Positive
		-MOI for absolute maternal indications (AMI)	Positive
		-C-section*** rates	Positive
		-Stillbirths among babies born by C-section	Positive
		-Very early neonatal death (<12 h) among babies born by C-section	Positive
		-Perception of the quality of care	Positive
Ridde (2003)	-Introduction of user fees in PHC facilities	-New curative consultations	Negative

*IUD: Intrauterine device. **TB: Tuberculosis. ***C-section: Cesarean section.

Introducing User Fees. We include ten studies reporting on the effects of introducing user fees, four out of which reporting mixed effects (Benjamin *et al.*, 2001; Chawla & Ellis, 2000; Jacobs & Price, 2004; Matee & Simon, 2000), three reporting negative effects (Kipp *et al.*, 2001; Kremer & Miguel, 2007; Ridde, 2003), two reporting positive effects (Audibert & Mathonnat, 2000; Richard *et al.*, 2007) and one reporting no effects (Mubyazi *et al.*, 2006).

In Papua New Guinea, Benjamin *et al.* (2001) reported that introducing user fees for obstetric services in four urban clinics and a general hospital was associated with an immediate decrease in attendance to obstetric care, followed by an increase in and stabilization of the frequency of attendances 12 months after introducing user fees. Despite quality improvements, introducing direct user fees and indirect insurance payments in government health facilities in Niger was associated with mixed effects on reporting an illness, no effects on seeking treatment and mixed effects on seeking formal treatment (Chawla & Ellis, 2000). In Cambodia, introducing user fees at a district referral hospital was associated with a decrease in admissions for dengue but no change in admissions for malaria, diarrhea/dysentery, respiratory infections, deliveries and other conditions (Jacobs & Price, 2004). More importantly, the study found that user fees had negative effects on hospital mortality rates and ability to pay. However, both total admissions and pediatric admissions remained unaffected by user fees. Matee & Simon (2000) reported that introducing user fees for dental health services provided by the government in Tanzania was associated with a decrease in dental attendance but no effect on either the demand for treatment or treatment pattern.

A decrease in utilization levels for outpatient services, deworming drugs and curative services was observed in Uganda, Kenya and Burkina Faso, respectively, after introducing user fees (Kipp *et al.*, 2001; Kremer & Miguel, 2007; Ridde, 2003). No effects were reported for introducing user fees in public health facilities in Tanzania on quality of care and malaria health-seeking behavior measured by malaria patient attendances (Mubyazi *et al.*, 2006).

Only two studies reported positive effects of introducing user fees. In Mauritania, the results of Audibert & Mathonnat (2000) were largely positive with respect to the improvement of the quality of care and the overall level of utilization of basic health facilities. Similarly in Burkina Faso, introducing user fees for emergency obstetric

care in an urban district hospital was associated with higher quality of care as well as higher emergency referrals from health centers, major obstetric interventions (MOI), MOI for absolute maternal indications (AMI) and Cesarean section (C-section) rates. User fees were also associated with lower stillbirths and lower very early neonatal death among babies born by C-section (Richard *et al.*, 2007). The findings of both studies suggest that users are willing to pay when the quality of health care improves. While Audibert & Mathonnat (2000) highlighted the importance of the supply of essential drugs and motivation of staff, Richard *et al.* (2007) highlighted the importance of the availability of equipment and the standardization of protocols.

Increasing User Fees. Five studies reported on the effects of increasing user fees, three out of which reporting negative effects on the majority of outcomes investigated (Bratt *et al.*, 2002; Cohen & Dupas, 2010; Issifou & Kremsner, 2004), one reporting mixed effects (Jacobs & Price, 2004) and one reporting no effects (Benjamin *et al.*, 2001).

In Ecuador, higher user fees for obstetric services in urban clinics and a general hospital was associated with a decrease in the utilization of these services measured by gynecology visits, intrauterine device (IUD) insertion visits, IUD revisits and ANC visits. Effect on seeking formal treatment, however, was mixed (Bratt *et al.*, 2002). A randomized malaria prevention experiment in Kenya also found that a higher price of antimalarial insecticide-treated bed nets (ITN) was associated with a decrease in demand (Cohen & Dupas, 2010). Similarly, Issifou & Kremsner (2004) found that an increase in consultation fees in a private hospital in Gabon had a negative effect on pediatric outpatient visits.

Jacobs & Price (2004) report mixed effects of higher user fees at a district referral hospital in Cambodia. While negative effects were observed with respect to admissions for malaria and dengue, positive effects were observed with respect to admissions for other conditions. Admissions for diarrhea/dysentery, respiratory infections and deliveries remained unaffected. More importantly, the study found that an increase in user fees had a negative effect on hospital mortality rates. However, no effects were observed with respect to ability to pay as well as admissions. No effects were reported by Benjamin *et al.* (2001) of higher user fees for delivery and postnatal care services in Papua New Guinea on institutional delivery.

Conclusion. The effect of introducing or increasing user fees in low- or middle-income countries is a controversial issue in the literature. The available evidence is limited and of low quality. The studies reviewed suggest, in general, that introducing or increasing user fees had a negative effect on the utilization of health services.

Although some of the studies reviewed reported that this effect was mitigated by quality improvements, the effects of simultaneously introducing user fees and quality improvement interventions remain unclear due to the absence or quality of existing evidence. Moreover, evidence on the longer-term effects of introducing or increasing user fees is lacking. Further research is needed to provide evidence in this regard.

4.3 ECONOMETRIC STRATEGY

Using the 2008 and 2014 waves of the Egypt Demographic and Health Survey (DHS), we employ the DD method to estimate the effects of introducing user fees on our health outcomes of interest. The use of DD is appropriate in our context as pre- and post-treatment health outcomes are observed for both accredited and contracted health facilities. Also as discussed in chapter 3, we could rely on DD to address our econometric concerns, most importantly endogeneity problems that could arise when comparing between heterogeneous facilities.

We replicate the estimation of equation (3.1) in chapter 3. In the replication, however, we adopt a new definition of treatment: introducing user fees. We also use two different waves of the Egypt DHS: 2008 and 2014. We include facilities that are "accredited only" in 2008 and continue to be so in 2014 in our control group. These are the facilities that, despite being accredited, did not contract with the FHF. The treatment group includes facilities that are accredited in 2008 and are both accredited and contracted in 2014. For each health facility i at time t, the model specification is as follows:

$$y_{it} = \alpha + \beta\, copay_{it} + \gamma\, d_{post} + \delta\, copay_{it} * d_{post} + \zeta\, fac_i + \eta\, dist_i + \varepsilon_{it} \quad (4.1)$$

The dependent variable y_{it} denotes a health outcome of interest y for facility i at time t. t = 0 for the baseline year (2008) and t = 1 for the follow-up year (2014). The

variable $copay_{it}$ = 1 if a facility is both accredited and contracted and $copay_{it}$ = 0 if a facility is accredited only. The coefficient β captures the differences between contracted and "accredited-only" facilities in 2008 prior to any contractual agreements with FHFs. The variable d_{post} = 1 for year 2014 and d_{post} = 0 for year 2008. The coefficient γ captures any time trends in health outcomes as well as other factors that could affect outcomes even if facilities did not contract with the FHF after being accredited. The parameter of interest is the DD estimator, δ, which is the coefficient of the interaction term $(copay_{it} * d_{post})$. The term equals one for contracted facilities in 2014 and zero otherwise. δ is the estimated change in the outcomes of contracted facilities relative to accredited facilities as a result of contracting with the FHF and, consequently, becoming entitled to collect user fees. The definitions of the facility-level controls fac_i, the district-level controls $dist_i$ and the error term ε_{it} are the same as discussed in model specification (3.1) in chapter 3.

For each health outcome, we report the results of estimating three specifications of equation (4.1). While the model's specification (1) includes no controls, specification (2) includes facility-level controls only and specification (3) includes both facility- and district-level controls. Reporting results from these three specifications enables us to check the extent to which the effects on outcomes are sensitive to variation in our control set. Standard errors obtained are robust across clusters defined by facility.

4.4 DATA

In this section, we discuss how we construct our dependent and explanatory variables. We use the 2008 and 2014 waves of the Egypt DHS to calculate our health outcomes of interest at the facility level. Information obtained from Egypt's MOH is used to capture interventions at the facility level and calculate facility-level controls. Data obtained from the Central Agency for Public Mobilization and Statistics (CAPMAS) is used to calculate a set of district-level social and economic controls. A number of regional dummies is included as well in our analyses whenever possible.

4.4.1 Dependent Variables

We make use of all the data made available by the Egypt DHS to construct indicators that reflect the utilization of family planning, ANC and delivery care services. Moreover, we construct an indicator of women's access to health care. These are the

outcomes that are expected to be affected by introducing user fees. The DHS data does not allow us to construct indicators of utilization of child care services. Alternatively, we construct an indicator of child mortality as a comprehensive measure of the health status of children.

To construct our health outcomes, we follow the same steps discussed in chapter 3. First, we spatially link women interviewed in each of the Egypt DHS waves to their nearest mapped facilities using Quantum GIS 2.8.2. Second, we recode and compute the health outcomes at the facility level using Stata 12.0. We, finally, combine all waves of the survey in a panel. Table 4.2 summarizes the description and sources of dependent variables included in the analyses of this study.

Table 4.2: Description and sources of dependent variables

	Outcome	Description*	Units	Source	Year(s)
Family planning	*mcp*	Proportion of women currently using any modern contraceptive method	Percent	Author's calculations based on Egypt DHS	2008, 2014
ANC	*ancprov*	Proportion of women attended for ANC by skilled health personnel	Percent	Author's calculations based on Egypt DHS	2008, 2014
	anc4	Proportion of women who received four or more ANC visits	Percent	Author's calculations based on Egypt DHS	2008, 2014
	anciron	Proportion of women who received iron supplements as an ANC component	Percent	Author's calculations based on Egypt DHS	2008, 2014
Delivery care	*delplac*	Proportion of live births delivered in a health facility	Percent	Author's calculations based on Egypt DHS	2008, 2014
	delassist	Proportion of live births whose delivery was assisted by skilled health personnel	Percent	Author's calculations based on Egypt DHS	2008, 2014
	delcaes	Proportion of live births delivered by C-section	Percent	Author's calculations based on Egypt DHS	2008, 2014
Access to care	*accmon*	Proportion of women with a "getting money for treatment" problem in accessing health care	Percent	Author's calculations based on Egypt DHS	2008, 2014
Child health status	*childmort*	Proportion of deaths at age 0–5 years to live-born children	Percent	Author's calculations based on Egypt DHS	2008, 2014

*Definitions are obtained from the World Health Organization (WHO).

Family Planning. We include an indicator that reflects current use of family planning methods: modern contraceptive prevalence *(mcp)*. *mcp* is a Millenium Development Goal (MDG) outcome indicator that tackles progress towards Target 5.B of MDG 5. We include *mcp* in our analyses as a proxy measure of access to reproductive health services. In this sense, higher *mcp* accelerates progress towards MDGs concerned with child mortality, HIV/AIDS and gender equality. The World Health Organization (WHO) defines *mcp* as "the percentage of women aged 15-49 years, married or in-union, who are currently using, or whose sexual partner is using, at least one method of modern contraception, regardless of the method used." According to the Guide to DHS Statistics, *mcp* is the quotient of the number of currently married women between ages 15 and 49 years who say they use one of the modern contraceptive methods[1] divided by the number of currently married women between ages 15 and 49 years, expressed as a percentage. As part of the Egypt DHS ever-married women (EMW) questionnaire, each female respondent is asked if she is currently doing something or using any method to delay or avoid getting pregnant and which method(s) she is using.

ANC. The first dimension of ANC considered is women's choice of the type of provider of ANC. We attempt to assess whether introducing user fees encourages or discourages women to shift from traditional birth attendants to doctors or trained nurses/midwives. We calculate the percentage of women with a birth in the last five years who received ANC by skilled health personnel *(ancprov)*. According to the Guide to DHS Statistics, *ancprov* is the quotient of number of women who were attended for ANC for their last birth and received care by skilled health personnel, divided by the number of women with a birth in the last five years, expressed as a percentage. The indicator reflects the use of skilled care.

The most important ANC outcome we include in our analyses, however, is ANC coverage (at least four visits) *(anc4)*. It is used as a global preferred indicator of access to and use of health care during pregnancy to track performance in maternal health programs. WHO defines *anc4* as "the percentage of women aged 15-49 with a live birth in a given time period who received ANC four or more times." According to the

[1] Female sterilization, male sterilization, the contraceptive pill, intrauterine contraceptive device, injectables, implants, female condom, male condom, diaphragm, contraceptive foam and contraceptive jelly, lactational amenorrhea method, country-specific modern methods and respondent mentioned other modern contraceptive methods.

Guide to DHS Statistics, *anc4* is the quotient of the numbers of women who received ANC for their last birth, according to grouped number of visits (four visits), divided by the number of women with a birth in the last five years, expressed as a percentage. As part of the Egypt DHS EMW questionnaire, each female respondent is asked how many times she received ANC during each of the pregnancies of her children born in the last five years. *anc4* is a MDG outcome indicator that tackles progress towards Target 5.B of MDG 5. WHO recommends that a woman receives at least four antenatal visits during a normal pregnancy to ensure that antenatal complications are detected and controlled at the earliest stage. A pregnant woman is expected to receive health interventions during antenatal visits that could be vital to her health and the health of her infant as well.

Despite the significance of *anc4*, the use of this sole indicator as a summary measure of access to and use of health care during pregnancy inappropriately emphasizes the number of visits at the expense of the content and process of care. Therefore, we include a third ANC outcome to reflect the utilization of nutrients during pregnancy: iron supplementation during pregnancy *(anciron)*. We calculate the percentage of mothers who received iron supplements during pregnancy. According to the Guide to DHS Statistics, *anciron* is the quotient of the number of women who received ANC for their last birth and who received iron supplements during pregnancy, divided by the number of women with a birth in the last five years who received ANC for their last birth, expressed as a percentage.

Delivery Care. Three indicators of utilization of delivery care services are included in our analyses: institutional delivery *(delplac)*, skilled assistance during delivery *(delassist)* and C-section delivery rates *(delcaes)*. The two former indicators are widely advocated for reducing maternal, perinatal and neonatal mortality. Through the indicator *delplac*, we attempt to investigate the effect of introducing user fees on access to childbirth facilities. *delplac* is also a proxy measure of maternal and neonatal morbidity and mortality. Women who give birth at a health facility are more likely to receive proper medical attention and care during delivery. Their infants as well are more likely to receive proper care after delivery. According to the Guide to DHS Statistics, *delplac* is the quotient of the numbers of live births whose deliveries took place in a health facility, divided by the number of live births in the last five years, expressed as a percentage.

The second but most important measure of delivery care we include in our analyses is *delassist.* WHO defines *delassist* as "the proportion of births attended by skilled health personnel." *delassist* is a MDG outcome indicator that tackles progress towards Target 5.A of MDG 5, which is to "reduce by three quarters, between 1990 and 2015, the maternal mortality ratio." Empirical literature provides evidence that wider access to professional care during pregnancy and childbirth reduces maternal mortality. Women assisted by skilled health personnel during delivery are less likely to die from any cause related to or aggravated by childbirth (Graham *et al.*, 2001). According to the Guide to DHS Statistics, *delassist* is the quotient of the number of live births assisted by medical provider (doctor or nurse/midwife) during delivery divided by the number of live births in the last five years, expressed as a percent. As part of the Egypt DHS EMW questionnaire, each female respondent is asked either a health professional (doctor or nurse/midwife) or other person (daya or other) or no one assisted with the delivery of each of her children.

Another measure of the availability of delivery care included in our analyses is the percentage of live births delivered by C-section *(delcaes). delcaes* is a key indicator of access to and use of health care during childbirth. We should be cautious, however, when interpreting this indicator as we recognize that very high levels of C-sections are as dangerous as very low levels. According to the Guide to DHS Statistics, *delcaes* is the quotient of the number of live births delivered by C-section divided by the number of live births in the last five years, expressed as a percent.

Access to Care. Besides antenatal and delivery care outcomes, we construct an indicator of women's access to health care. We calculate the percentage of women with a "getting money for treatment" problem in accessing health care for themselves *(accmon).* This indicator investigates whether introducing user fees deepens the role of money as an impediment to women to access health care. *accmon* is calculated as the quotient of the numbers of women who reported getting money for treatment as a problem in accessing health care for themselves, divided by the number of interviewed women, expressed as a percentage.

Child Health Status. We finally calculate an indicator of child mortality as a comprehensive measure of the health status of children: under-five mortality rate *(childmort). childmort* is a MDG impact indicator that tackles progress towards Target

4.A of MDG 4, which is to "reduce by two-thirds, between 1990 and 2015, the under-five mortality rate." We calculate *childmort* as the quotient of the number of deaths at age 0–5 years divided by the number of live-born children in the last five years, expressed as a percentage.

4.4.2 Explanatory Variables

The main explanatory variable in our model is $copay_{it}$. The variable draws on information from Egypt's MOH regarding whether and when a facility is unreformed, accredited or contracted. After 2008, the main difference between accredited and contracted facilities is that the latter are entitled to collect user fees from beneficiaries. To capture the effects of user fees in this context, we define treatment as having a contractual agreement with the relevant FHF by 2014 after being "accredited only" in 2008. A facility is considered to be control if it is accredited in 2008 and continues to be so in 2014. Unreformed facilities as well as facilities that are originally contracted in 2008 are removed from our dataset.

In all regression models, we include facility-level characteristics, district-level social and economic characteristics as well as regional dummies to control for any potential discrepancies at facility, district and regional levels, respectively. A detailed list of all the control variables included in our analyses is provided in chapter 3 (see Table 3.4).

4.5 RESULTS

In this section, we present the descriptive statistics of the subsample used to estimate the effect of introducing user fees and discuss the estimated effect of introducing user fees on family planning and maternal and child health.

4.5.1 Descriptive Statistics

The descriptive statistics of the facility-level characteristics, the district-level characteristics and the health outcomes of our subsample are reported in Table 4.3. Some disparities are observed between health facilities with respect to characteristics and outcomes. In parallel, we observe that Egypt performs well with respect to key antenatal and delivery care outcomes. However, Egypt's cesarean section (C-section) delivery rate is way above the ideal rate set by the international healthcare community to be between 10-15 percent (World Health Organization, 2015).

Table 4.3: Descriptive statistics

Variable	N	Mean	Standard deviation	Minimum	Maximum
Facility characteristics					
Practitioners	304	5.322	7.077	0.000	57.000
Specialists	270	0.822	1.767	0.000	13.000
Pharmacists	298	5.872	6.310	0.000	36.000
Nurses	302	14.146	9.760	1.000	60.000
Lab technicians	296	1.149	1.256	0.000	7.000
X-ray technicians	268	0.216	0.604	0.000	3.000
Health observers	298	1.336	1.042	0.000	5.000
Social workers	270	1.030	1.478	0.000	7.000
Building condition	268	1.575	0.686	0.000	2.000
Population coverage	308	38.528	61.404	1.866	465.000
District characteristics					
Illiteracy	328	29.059	10.916	4.980	51.210
Unemployment	328	9.772	3.741	2.150	23.200
Income dependency	328	3.806	12.302	0.290	55.982
Inaccessibility to electricity	328	0.737	0.782	0.023	8.871
Inaccessibility to potable water	328	2.595	4.602	0.012	27.134
Family size	328	4.192	0.368	3.310	5.240
Household (HH) overcrowding	328	1.150	0.099	0.860	1.510
Population size	328	32.732	18.896	3.566	117.380
Health outcomes					
Family Planning					
Modern contraceptive prevalence	166	53.902	16.291	0.000	91.667
ANC					
ANC by skilled health personnel	136	84.691	17.832	20.000	100.000
4+ visits	135	79.377	19.911	20.000	100.000
Iron supplementation	136	60.069	23.754	0.000	100.000
Delivery care					
Institutional delivery	136	85.340	19.180	12.500	100.000
Skilled-assisted delivery	136	89.881	16.848	12.500	100.000
C-section delivery	136	47.261	25.742	0.000	100.000
Access to care					
Money barrier reported	166	21.438	24.465	0.000	92.857
Chilh health status					
Under-5 mortality	146	2.513	4.937	0.000	30.000

N denotes the number of observations.

Tables 4.4 presents differences in facility-level characteristics between contracted and "accredited only" health facilities. We use the two-sample t-test to check whether the means of the two groups differ significantly. On average, we find that accredited facilities are better than contracted facilities with respect to more than half of the labor force characteristics. However, the differences observed are not significant in seven out of eight labor force indicators. The only significant difference is observed with respect to the number of social workers (*socwork*) in facilities. In parallel, we do not observe significant differences in the building condition (*inf*) and population coverage (*pop*) between the two groups of facilities (Tables 4.4).

Table 4.4: Two-sample t-test of facility characteristics of accredited and contracted facilities

	Accredited	Contracted	Difference
Practitioners	5.126	6.320	-1.194
			(1.553)
Specialists	0.843	0.700	0.143
			(0.430)
Pharmacists	6.064	4.875	1.189
			(1.410)
Nurses	13.921	15.333	-1.412
			(2.180)
Lab technicians	1.210	0.833	0.376
			(0.280)
X-ray technicians	0.246	0.050	0.196
			(0.146)
Health observers	1.355	1.240	0.115
			(0.229)
Social workers	1.148	0.350	0.798**
			(0.353)
Building condition	1.536	1.773	-0.237
			(0.159)
Population coverage	39.297	34.560	4.737
			(13.479)

Standard errors are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

Similarly, we use the two-sample t-test to check whether district-level characteristics of contracted and "accredited only" facilities differ significantly. The results of our tests are reported in Table 4.5. We observe that contracted facilities are located in districts with more favorable socio-economic profiles compared to districts where accredited facilities are located.

Table 4.5: Two-sample t-test of district characteristics of accredited and contracted facilities

	Accredited	Contracted	Difference
Illiteracy	29.081	28.944	0.136
			(2.345)
Unemployment	10.139	7.826	2.313***
			(0.783)
Income dependency	4.418	0.556	3.863
			(2.625)
Inaccessibility to electricity	0.775	0.537	0.237
			(0.167)
Inaccessibility to potable water	2.682	2.132	0.550
			(0.988)
Family size	4.200	4.146	0.054
			(0.079)
HH overcrowding	1.157	1.110	0.047**
			(0.021)
Population size	31.818	37.579	-5.761
			(4.033)

Standard errors are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

However, significant differences between contracted and accredited facilities are only observed with respect to unemployment (*unemp*) and HH overcrowding (*crowd*).

4.5.2 Estimated Effects of Introducing User Fees

Table 4.6 indicates that introducing user fees does not have a significant effect on modern contraceptive prevalence (*mcp*), a proxy measure of access to reproductive health services. A possible explanation is that even if a woman has access to a contracted facility, she would still have to pay a highly subsidized price to obtain a family planning method. This price is equal to the price incurred by a woman with access to an "accredited only" facility. In this regard, we note that Egypt's MOH offers family planning services to all women at nominal fees in an effort to slow down the rapid population growth.

In parallel, Table 4.6 indicates that having access to a contracted facility that introduced user fees is associated with a higher likelihood of receiving ANC by skilled health personnel (*ancprov*) as well as a higher likelihood of receiving at least four ANC visits (*anc4*). As discussed earlier, *ancprov* and *anc4* reflect the access to and use of skilled care during pregnancy. We find that both *ancprov* and *anc4* increased significantly by 18 ppts between 2008 and 2014 among women with access to contracted facilities compared to women with access to "accredited only" facilities.

Table 4.6: Estimated effects of introducing user fees

	Outcome	DD		
		(1)	(2)	(3)
Family planning	Modern contraceptive prevalence	-3.809	-0.810	-2.943
		(5.418)	(7.630)	(7.129)
ANC	ANC by skilled health personnel	16.112**	15.153*	17.929**
		(6.890)	(8.545)	(8.795)
	4+ visits	16.935**	13.154	17.776*
		(7.564)	(9.060)	(10.008)
	Iron supplementation	24.118***	22.408**	23.589**
		(7.959)	(9.894)	(8.978)
Delivery care	Institutional delivery	-5.409	1.763	8.847
		(6.302)	(8.583)	(8.701)
	Skilled-assisted delivery	-4.608	2.851	7.818
		(5.837)	(7.348)	(7.226)
	C-section delivery	6.101	14.296	16.104
		(8.140)	(10.704)	(11.902)
Access to care	Money barrier reported	9.534	6.966	8.947
		(10.074)	(12.426)	(12.687)
Child health status	Under-5 mortality	2.835**	2.890	1.602
		(1.374)	(1.980)	(1.946)

Each row represents a separate regression. Robust standard errors across clusters are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

Moreover, having access to a contracted facility is associated with a higher likelihood of receiving iron supplements during pregnancy *(anciron)*. Table 4.6 shows that the proportion of women with access to contracted facilities, who receive iron supplements during pregnancy *(anciron)*, increased significantly by 24 ppts between 2008 and 2014 compared to women with access to "accredited only" facilities. Thus, while there is a general trend in Egypt towards increased use of ANC services, contracting can add to this trend.

The results reported with respect ANC outcomes is inconsistent with the theory of consumer demand. The positive effects captured in contracted facilities despite introducing user fees could, however, be justified by the higher quality of services provided by these facilities compared to "accredited only" facilities. The FHF contracts with facilities to deliver a package of services to the population. As the contracting agency, the FHF sets and supervises the rules and eligibility criteria for these facilities. The FHF has the right to contract with a facility only if it applies a set of guidelines as discussed in chapter 2. For instance, only a facility that is prepared to adopt the Family Health Model (FHM) can contract with the FHF. Thus, the facility should fulfill the family medicine accreditation requirements to be eligible to contract. Additionally, the facility should possess a standard catalogue of equipment. This

catalogue was originally developed to ensure that facilities have the equipment necessary to deliver the BBP of services defined under the HSRP. Hence, contracted facilities are initially better than "accredited only" facilities.

In this regard, it is also important that contracted facilities are subject to additional supervision by the FHF. FHFs have an autonomous supervision system that focuses on the administrative and financial arrangements of contracted facilities. While all health facilities are regularly visited by supervisors from the Health District, contracted facilities also receive tri-monthly visits from the FHF.

Grun & Ayala (2006) supports our hypothesis that contracted facilities are expected to provide higher quality of services. The study provides evidence that contracting is associated with a positive effect on accreditation scores. According to Grun & Ayala (2006), the accreditation score of a facility increases by about 3 ppts right after it contracts with the relevant FHF. As discussed in chapter 3, a higher accreditation score reflects higher compliance with standards in eight categories, each includes several measures of key processes, activities and outcomes that facilities should achieve. In this sense, the findings of Grun & Ayala (2006) provide additional justification of the positive effects captured with respect to ANC outcomes in contracted facilities despite introducing user fees in these facilities.

With respect to delivery care, Table 4.6 shows that introducing user fees is associated with positive but insignificant effects on institutional delivery (*delplac*) and skilled assistance during delivery (*delassist*). While *delplac* reflects access to childbirth facilities, *delassist* reflects access to and use of professional care during pregnancy and childbirth. A higher quality of care is expected to drive utilization up and introducing user fees is expected to drive utilization down. Our findings on delivery care suggest that the negative effects associated with introducing user fees weaken the positive effects associated with providing a higher quality of care. In parallel, Table 4.6 reports on C-section delivery rates (*delcaes*) that is also a key indicator of access to and use of health care during childbirth. We find that introducing user fees does not have a significant effect on *delcaes*. It is important to note that we initially expected that introducing user fees would be associated with a significant decrease in *delcaes* as Egypt's C-section delivery rate is far higher than the ideal. Our initial expectation is supported by the fact that some countries attempt to reduce the demand

for elective C-section by introducing a copayment when C-section is not medically indicated (Chen *et al.*, 2014).

In addition, Table 4.6 shows that having access to a contracted facility that introduced user fees is not associated with a significant change in the likelihood of reporting money as an impediment to access health care *(accmon)*. This finding suggests that introducing user fees would not restrict women's access to care if accompanied by quality improvement and a pro-poor exemption policy. We expect that women have become more aware of the exemption policy in place since 2003. As noted earlier in chapter 3, Ministerial Decree 147/2003 that institutionalizes user fees includes an exemption clause for patients who cannot afford to pay. Another possible justification is that services provided by contracted facilities became relatively more appealing to the population on two grounds. First, we realize that private healthcare providers in Egypt had raised their fees significantly during the period 2008-2014. Consequently, user fees incurred at public contracted facilities have become relatively less costly compared to the fees set by private providers. Second, contracted facilities stand a better chance than "accredited only" facilities of competing with private providers as contracted facilities offer a higher quality of care.

Finally, we find that introducing user fees does not have a significant effect on under-five mortality (*childmort*) during the period 2008-2014 (Table 4.6).

4.6 ROBUSTNESS CHECKS

To support our main results reported in the previous section, we re-estimate equation (4.1) using two new definitions of treatment and drawing on two new subsamples. In addition, we test the validity of the parallel-trend assumption and run a number of placebo tests to ensure that our DD model is correctly specified.

4.6.1 Supporting Estimations

To support our results reported in Table 4.6, we re-estimate equation (4.1) using two new definitions of treatment and drawing on two new subsamples. Our aim is to capture the effect of introducing user fees by comparing between changes in the outcomes of unreformed facilities after accreditation versus changes in the outcomes of unreformed facilities after both accreditation and user fee introduction. We could interpret this effect as the incremental effect of introducing user fees in accredited

facilities. We only include observations of facilities that are observed in both years 2008 and 2014. To capture the effect of "accreditation only", we use a subsample that includes health facilities that are unreformed in both 2008 and 2014 (control group) and facilities that are unreformed in 2008 and are "accredited only" in 2014 (treatment group). To capture the effects of combining accreditation and contracting, we use a subsample that includes facilities that are unreformed in both 2008 and 2014 (control group) and facilities that are unreformed in 2008 and contracted in 2014 (treatment group). We include all controls in this analysis.

Our re-estimation results provide evidence that although contracting implies introducing user fees, it could mitigate the negative effects associated with accreditation (Table 4.7). This supports the findings of our main results that introducing user fees could even be associated with positive effects.

Table 4.7: Estimated effect of accreditation versus contracting

	Outcome	DD	
		Accredited vs unreformed	Contracted vs unreformed
Family planning	Modern contraceptive prevalence	4.581	8.620
		(3.595)	(6.699)
ANC	ANC by skilled health personnel	-7.846*	4.391
		(4.095)	(8.809)
	4+ visits	-9.572**	1.263
		(4.586)	(9.236)
	Iron supplementation	3.606	-0.798
		(7.145)	(12.414)
Delivery care	Institutional delivery	-2.930	2.586
		(4.257)	(10.678)
	Skilled-assisted delivery	-3.508	-2.344
		(3.626)	(6.175)
	C-section delivery	-6.752	-3.552
		(4.825)	(11.095)
Access to care	Money barrier reported	-2.717	-3.243
		(4.063)	(6.493)
Child health status	Under-5 mortality	1.570	-0.735
		(1.313)	(2.088)

Each row represents a separate regression. Robust standard errors across clusters are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

4.6.2 Parallel Trends

The key assumption of DD is parallel trends. We must verify that differences in health outcomes between treated and control facilities would have remained the same in the absence of treatment. However, it is not feasible to test the validity of parallel trends for our main results reported in Table 4.6 due to data limitations. As previously noted, to obtain our main results, facilities that are "accredited only" in 2008 and are both accredited and contracted in 2014 are included in the treatment group. The control group includes facilities that are "accredited only" in both 2008 and 2014. While we observe facilities that could be defined as control in 2005, we do not observe any facility in 2005 that could be defined as treated. We have observations for facilities that are accredited in 2005, 2008 and 2014, but we do not observe facilities that are accredited in 2005 and 2008 and became contracted by year 2014.

Alternatively, we test for parallel trends to verify the results reported in Table 4.7. These are the results used to support our main results reported in Table 4.6. Verifying the robustness of our supporting results is the second-best option given the limitation in our data. First, we test for the robustness of the estimates reported in the third column of Table 4.7. These are the estimates that capture the effect of accreditation. We use the data of the study period 2005-2008 to evaluate the validity of parallel trends during the period 2008-2014. We only keep observations of facilities that are unreformed in both 2005 and 2008 and that are later observed in 2014 as either unreformed or accredited. We only include observations of facilities that are observed in all three years. We drop observations of facilities that are either accredited or contracted in 2005. For the period 2005-2008, we define facilities that became accredited after 2008 as treated and facilities that remain unreformed after 2008 as control.

We provide visual evidence on the absence of an existing trend that could invalidate the DD assumption made to estimate the effects of accreditation. The annual mean proportion of all our health outcomes are plotted in Figure 4.1. We find that the parallel-trend assumption is broadly satisfied for the majority of outcomes.

Figure 4.1: Parallel trends in outcomes

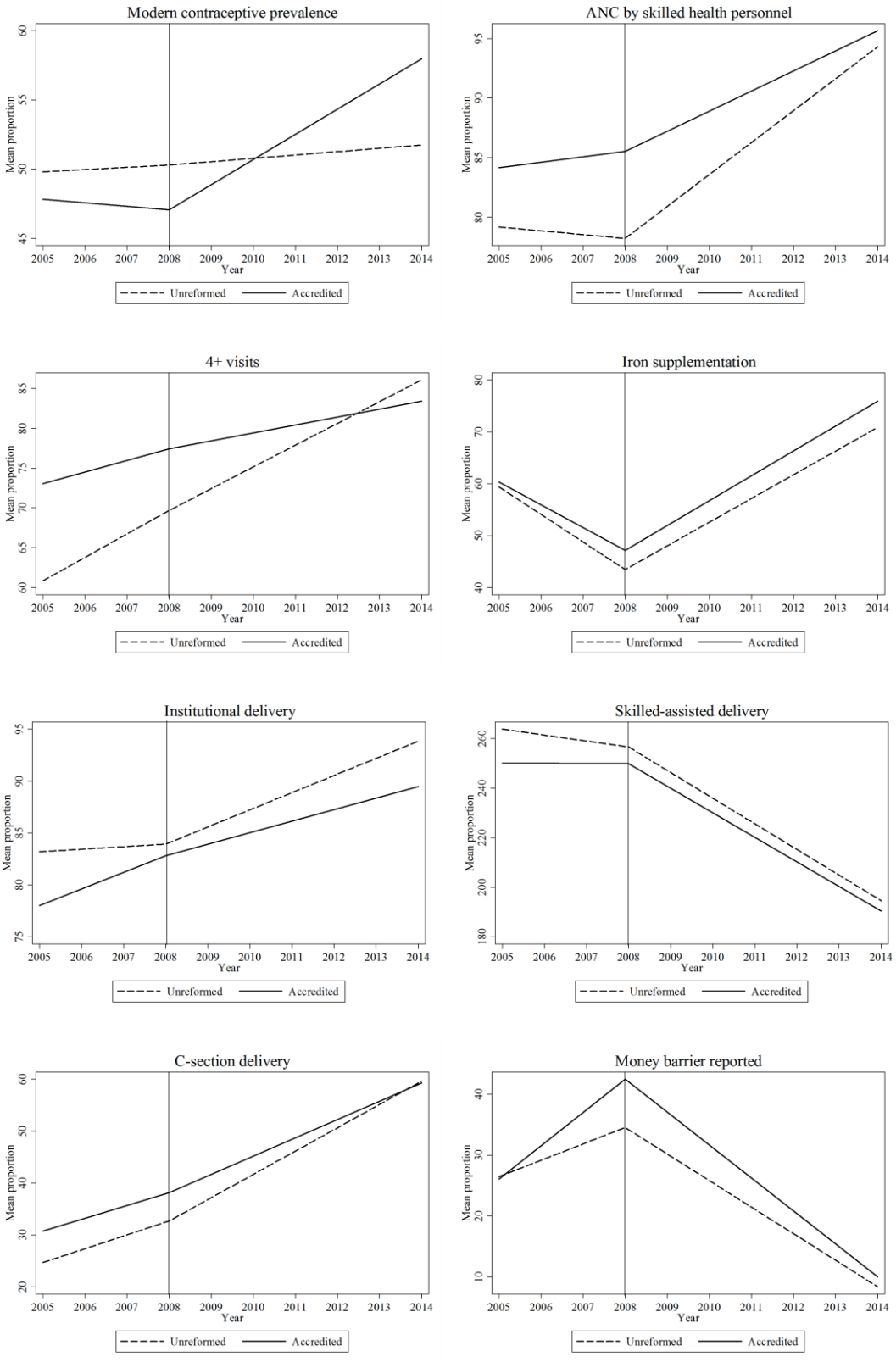
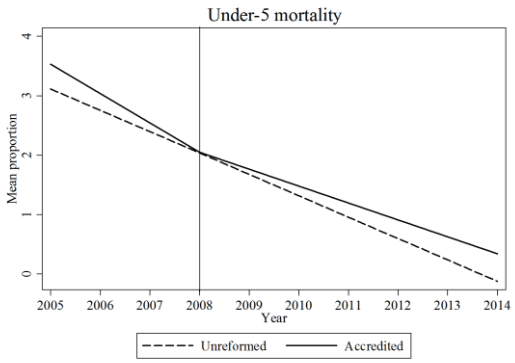


Figure 4.1: Parallel trends in outcomes (cont'd)



For some outcomes, however, we should initially match treated and control facilities. The inclusion of facility controls leads to a stricter satisfaction of the parallel-trend assumption for *mcp*, *anc4*, *delplac* and *childmort*. The inclusion of both facility and district controls leads to a stricter satisfaction for *delassist*.

Second, we test for the robustness of the estimates reported in the fourth column of Table 4.7. These are the estimates that capture the effect of contracting. We only include observations of facilities that are unreformed in both 2005 and 2008 and that either remain to be so or became contracted by year 2014. Observations of facilities that are accredited or contracted in 2005 are dropped. For the period 2005-2008, a facility is considered treated if it became contracted after 2008. Facilities that remain unreformed after 2008 belong to the control group.

The annual mean proportion of health outcomes are plotted in Figure 4.2. We find that the parallel-trend assumption is broadly satisfied for our outcomes of interest.

4.6.3 Placebo Test

We follow the placebo test of Bertrand *et al.* (2004) to confirm the robustness of the results reported in Table 4.7. We define "false" lagged accreditation and contracting interventions to inspect our health outcomes before the interventions. The pre-treatment estimations would yield null results if our DD model is correctly specified. We use the data of the period 2005-2008 to verify the results of our study period 2008-2014. We use the same data and definitions of treatment used earlier to test for the parallel-trend assumption. All controls are included in both placebo tests.

Figure 4.2: Parallel trends in outcomes

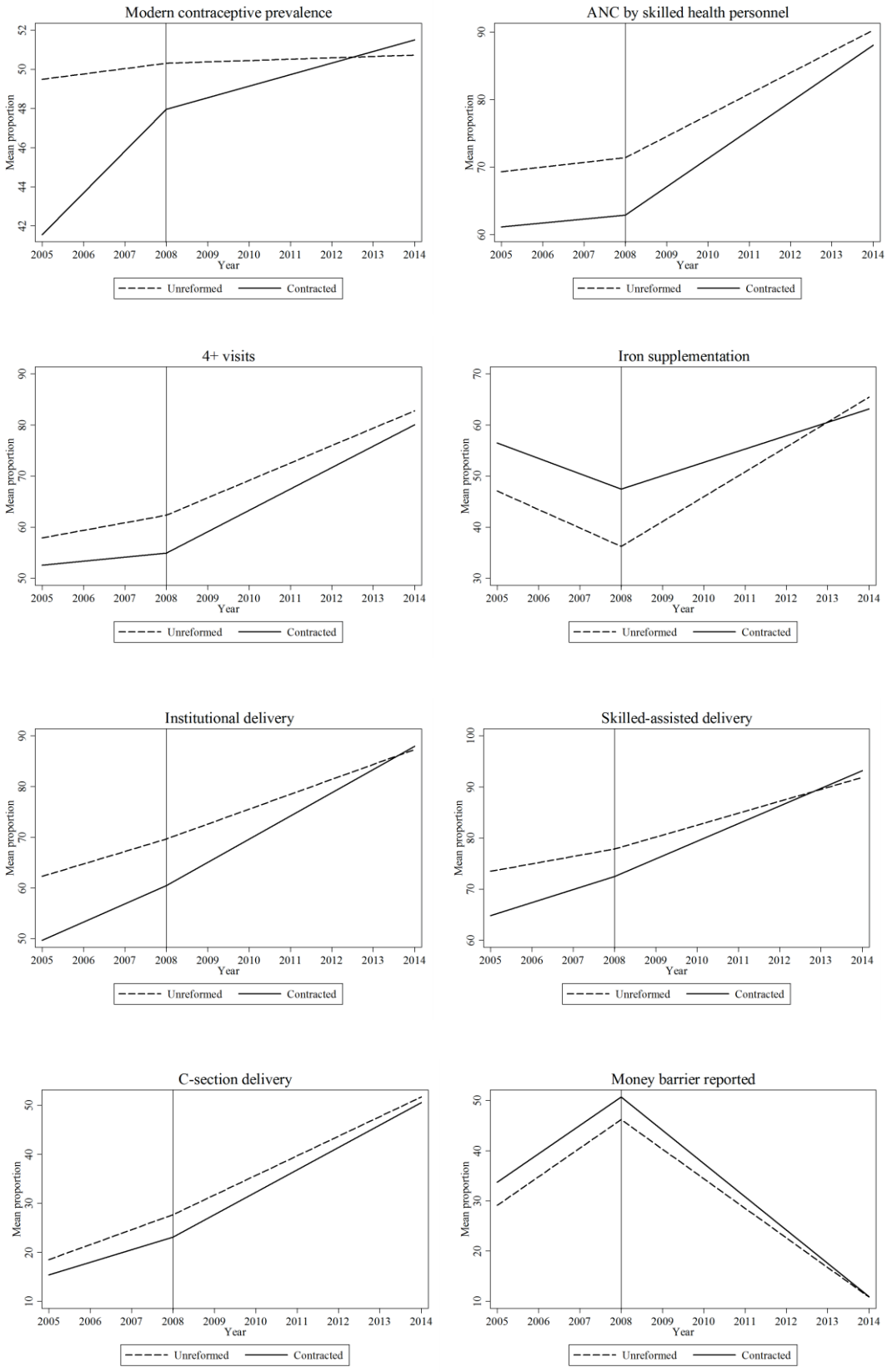
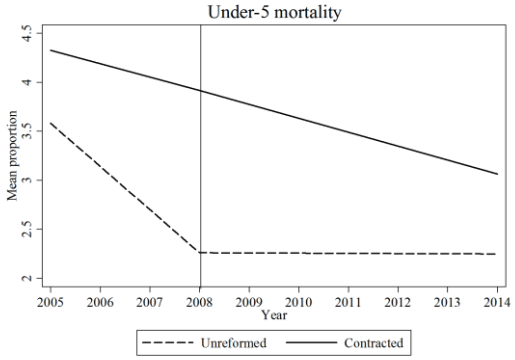


Figure 4.2: Parallel trends in outcomes (cont'd)



The results of the placebo test of accreditation are reported in the third column of Table 4.8. The estimates of all outcomes are not significantly different from zero. That is, differences between the outcomes of accredited and unreformed facilities reported in the third column of Table 4.7 only emerged after accreditation. This finding provides further evidence on parallel trends. Our treatment causes the effects, rather than the other way around.

Table 4.8: Estimated effects of placebo accreditation and contracting

	Outcome	DD	
		Accredited vs unreformed	Contracted vs unreformed
Family planning	Modern contraceptive prevalence	-1.260	4.477
		(4.492)	(3.992)
ANC	ANC by skilled health personnel	-5.477	2.119
		(7.745)	(6.904)
	4+ visits	-4.478	0.747
		(8.262)	(6.393)
	Iron supplementation	1.487	1.982
		(8.860)	(6.843)
Delivery care	Institutional delivery	4.057	2.876
		(6.971)	(6.628)
	Skilled-assisted delivery	6.960	4.727
		(7.009)	(6.158)
	C-section delivery	-3.443	-1.052
		(8.983)	(4.980)
Access to care	Money barrier reported	9.781	-3.689
		(8.003)	(6.084)
Child health status	Under-5 mortality	-0.404	1.093
		(2.791)	(1.786)

Each row represents a separate regression. Robust standard errors across clusters are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

The results of the placebo test of contracting are reported in the fourth column of Table 4.8. Similarly, the estimates of all outcomes are not significantly different from zero.

4.7 CONCLUSION

In this chapter, we use data from two waves of the Egypt DHS (2008 and 2014) to investigate the medium-term effect of introducing user fees on the utilization of family planning, ANC and delivery care services, women's access to health care, and child health status.

Using DD, we find that introducing user fees has no effect on the utilization of family planning and delivery care services. We also find that user fees do not hinder women's access to care. No effect is found on child health status as well. The lack of effect has been attributed to the fact that higher utilization of care, probably due to the expected quality improvement, did not offset the decrease in utilization due to introducing user fees.

The only positive effects observed of introducing user fees are with respect to the utilization of ANC services. Having access to a contracted facility that introduced user fees is associated with a higher likelihood of receiving ANC by skilled health personnel, a higher likelihood of receiving at least four ANC visits as well as a higher likelihood of receiving iron supplements during pregnancy. These positive effects captured in contracted facilities could be justified by the higher quality of ANC services provided by these facilities compared to "accredited only" facilities.

Our findings suggest that, in general, user fees are ineffective as a stand-alone policy. Introducing user fees should be part of a broader package of interventions that include addressing the quality of care in order to offset reduction in care utilization.

5. THE EFFECT OF INTRODUCING USER FEES TOGETHER WITH QUALITY IMPROVEMENT INTERVENTIONS ON FAMILY PLANNING AND MATERNAL AND CHILD HEALTH

5.1 INTRODUCTION

A Health Sector Reform Program (HSRP) was launched in Egypt at the national level in 1997. The main aim of the program is to address persistent needs in primary health care (PHC) in general and in maternal and child health in particular. Contracted health facilities that reached the highest phase of reform are subject to a broader package of interventions compared to unreformed or even "accredited only" facilities. On the supply side, interventions target quality improvement through facility accreditation and performance-based financing (PBF) of healthcare providers. As discussed in chapter 3, a facility accreditation program was introduced as an organized process to monitor the quality of services and ensure compliance with quality standards. In parallel, a PBF scheme was introduced in contracted facilities, according to which, monthly incentives are paid to health providers, who deliver a basic benefit package (BBP) of health services, based on pre-defined performance criteria (see chapter 2). On the demand side, user fees were introduced in contracted facilities. Uninsured beneficiaries became required to pay registration and renewal fees in addition to copayment fees that include visit fees, drug copayment and copayments for other interventions (see chapter 4).

A review of the available evidence on user fees suggests that user fee introduction can be associated with negative effects on the utilization of health services, although quality improvement may help maintain utilization in some cases (Lagarde & Palmer, 2008). However, the effectiveness of simultaneously introducing user fees and quality improvement interventions in low- and middle-income settings remains unclear due to the absence or quality of existing evidence. In a study of demand effects in Niger, introducing user fees was associated with mixed effects on reporting an illness, no effects on seeking treatment and mixed effects on seeking formal treatment despite quality improvements (Chawla & Ellis, 2000). In contrast, the findings of two studies

in Mauritania and Burkina Faso suggest that users are willing to pay when the quality of health care improves (Audibert & Mathonnat, 2000; Richard *et al.*, 2007).

This chapter contributes to the available evidence and adds to the results obtained in chapter 4 by allowing us to estimate the net effect of the combination of user fees and quality improvement instead of estimating the effect of user fees as a stand-alone policy. Combining difference-in-differences (DD) with propensity score matching (PSM), we estimate the effect on the utilization of family planning, antenatal care (ANC) and delivery care services, women's access to health care, and child health status during the period 2000-2008.

The chapter proceeds as follows. Section 5.2 provides background information for our study; section 5.3 discusses the econometric strategies used in this chapter; section 5.4 constructs our dependent and explanatory variables; section 5.5 presents the descriptive statistics together with the main results of our estimations; section 5.6 assesses the robustness of our estimation results; and section 5.7 concludes.

5.2 BACKGROUND

In this section, we describe the interventions in question, discuss the anticipated effect of combining these interventions and review evidence on their effectiveness in a low- or middle-income settings.

Cost Sharing and Quality Improvements under the HSRP. As discussed in chapter 4, user fees were introduced in 2003 to increase the ability of the FHF to generate revenues. As determined by Ministerial decree 147/2003, uninsured beneficiaries are required to pay registration fees, fees for renewal of registration as well as copayment fees that include visit fees, drug copayment and copayments for other interventions in contracted health facilities in all governorates.

However, user fee introduction was preceded by two main interventions to improve the quality of health care. First, a facility accreditation program was introduced under the HSRP in 2000. The main aim of facility accreditation is to improve the quality of PHC services provided by facilities participating in the HSRP. For a facility to get accredited, it must comply with pre-determined nationally established accreditation standards. As noted in chapter 3, optimal standards in eight categories were developed to assess key processes, activities, or outcomes that facilities should achieve. These

categories are patient rights, patient care, safety, management of support services, management of information, quality improvement program, family practice and management of the facility.

Second, a PBF scheme was integrated in the HSRP in 2001. According to this scheme, financial incentives were provided to contracted healthcare providers on a monthly basis based on pre-defined performance criteria. The percentage of incentives was determined based on the monthly performance of the providers, which is assessed through a set of 11 indicators: number of visits per day per physician, number of drugs per visit, rate of patient referral to the district hospital, rate of completion of visit encounter forms, patient satisfaction rate, rate of completion of medical records data, years of protection provided by contraceptive methods, number of children fully vaccinated in the catchment area, patient waiting time, number of ANC visits per pregnant woman and adherence to medical protocols. These indicators cover all aspects of service provision, whether curative or preventive, and maintains efficiency and quality. Each has a weight and standard.[1]

Anticipated Effect of Introducing User Fees and Quality Improvements. The simultaneous introduction of user fees and quality improvements is expected to drive demand in two opposite directions. As discussed in chapter 4, introducing user fees as a stand-one policy is expected to drive health service utilization down as the price incurred by the consumer increases. However, quality improvements are expected to drive utilization up due to the increased willingness to pay for an improved quality of service. The net effect of combining both interventions depends mainly on whether the positive effect of quality improvement outweighs the negative effect of user fee introduction. The net effect can also be affected by whether the user fees collected are used by the collector facility to improve the quality of care, and if so, what proportion of fees is used. Moreover, the net effect can be affected by whether an exemption policy is in place, and if so, how effective this policy is. We highlight the importance of outreach activities as well in mitigating the negative effect on access to health care that associates user fee introduction. Higher utilization can be associated with innovative outreach activities of health facilities.

[1] A detailed description of the PBF scheme is provided later in chapter 6.

Evidence on the Effect of Introducing User Fees and Quality Improvements. The available evidence on the net effect of introducing user fees accompanied by quality improvement interventions is limited. A review of the studies reviewed in chapter 4 suggests that even the available evidence is not conclusive, where two studies reported positive effects (Audibert & Mathonnat, 2000; Richard *et al.*, 2007), two studies reported mixed effects (Chawla & Ellis, 2000; Jacobs & Price, 2004) and one study reported no effects (Mubyazi *et al.*, 2006).

In Mauritania, the results reported by Audibert & Mathonnat (2000) of introducing user fees in all health facilities were largely positive with respect to improvement in the quality of health care and the overall level of utilization of basic health facilities. Quality improvements included higher supply of essential drugs and better-motivated staff. These results provide evidence that beneficiaries are willing to pay when the quality of health care improves. Similar positive effects were observed in Burkina Faso, where Richard *et al.* (2007) reported that introducing user fees for emergency obstetric care in an urban district hospital was associated with increased referrals from health centers and increased major obstetric interventions (MOI). These positive effects were probably driven by the improvement in quality of care due to the availability of all items required for the management of severe maternal conditions and to the standardization of the protocols.

However, Chawla & Ellis (2000) found that introducing direct user fees and indirect insurance payments in government health facilities in Niger was associated with mixed effects on reporting an illness, no effects on seeking treatment and mixed effects on seeking formal treatment. This mixed evidence was reported despite quality improvement interventions that included improving drug availability, training health personnel to use standard diagnosis and treatment protocols, strengthening management capacity, as well as improving supervisory and managerial capacity. Similar mixed effects were observed by Jacobs & Price (2004) in Cambodia, where increasing user fees at a district referral hospital had negative effects on admissions for malaria and dengue, positive effects on admissions for other conditions, and no effects on admissions for diarrhea/dysentery, respiratory infections, deliveries and total admissions. Prior to increasing user fees, quality improvement interventions at the hospital included strengthening the managerial capacity of its staff, and their diagnosis and treatment competence.

The findings of Mubyazi *et al.* (2006) support the evidence that introducing user fees would not have positive effects, especially when fees are not associated with a simultaneous improvement in the quality of services. The study reported no effects of introducing user fees in public health facilities in Tanzania on quality of care and malaria health-seeking behavior measured by malaria patient attendances.

In conclusion, evidence on the effect of introducing user fees accompanied by quality improvement interventions in low- or middle-income countries is limited, mixed and of low quality. There is some evidence that introducing user fees can increase service utilization if they are associated by substantial and sustained improvement in the quality of care. However, further investigation is required before a generalization can be made.

5.3 ECONOMETRIC STRATEGY

In this section, we discuss the two methods used to estimate the effects of introducing user fees and quality improvement under the HSRP. We start with considering DD. Complementarily, the DD approach is combined with PSM. To do so, we first match treated and control facilities based on pre-treatment characteristics. We then estimate the effects of treatment using DD.

DD. To remove potential biases, we use DD to estimate the effect of introducing user fees and quality improvements. The DD setup is also appropriate in our context as health outcomes for both contracted and non-contracted health facilities are observed both at the baseline and post treatment. We re-estimate the DD specification used in chapter 3 but using a new definition of treatment: introducing of user fees accompanied by quality improvements. Three different waves of the Demographic and Health Survey (DHS) are used: 2000, 2005 and 2008. Health facilities that are non-contracted in 2000 and continue to be so are included in the control group. Facilities that contracted with the FHF by 2005 or 2008 are included in the treatment group.

For each health facility i at time t, the model specification is as follows:

$$y_{it} = \alpha + \beta\, cont_{it} + \gamma\, d_{post} + \delta\, cont_{it} * d_{post} + \zeta\, fac_i + \eta\, dist_i + \varepsilon_{it} \qquad (5.1)$$

where y_{it} denotes health outcomes of interest; $cont_{it}$ is a treatment dummy variable that switches on if a facility is contracted; d_{post} is a time-period dummy variable that switches on in the follow-up year (2005 or 2008); the interaction term $(cont_{it} * d_{post})$ captures treatment after the baseline year 2000; δ is our coefficient of interest that captures the effect of contracting on each respective outcome at the facility level; fac_i is a vector of facility-level characteristics; $dist_i$ is a vector of district-level characteristics and ε_{it} is the error term that captures the effect of any unobservables.

Similar to chapter 3, the results of estimating three specifications of equation (5.1) are reported for two periods: 2000-2005 and 2005-2008. The first model's specification includes no controls; the second includes facility-level controls only and the third included both facility and district-level controls. The bootstrap method is used to estimate the standard errors in all regressions.

DD PSM. As noted in chapter 3, endogeneity is our main concern when estimating the effects of introducing user fees and quality improvements. Facilities are likely to differ along observable characteristics (e.g., socio-economic background) and non-observable characteristics (e.g., managerial competence). These potential differences require proper treatment of self-selection into treatment. Treating other potential sources of endogeneity such as omitted-variable bias is also required. To do so, we follow the DD Kernel PSM method discussed in chapter 3.

A facility is considered to be treated if it has a contractual agreement with the FHF in 2005 or 2008. A facility is considered to be control if it continues to be non-contracted in 2005 or 2008. First, treated and control districts are matched based on pre-treatment observable characteristics, which are captured by the eight indicators of the socio-economic vulnerability index used by the GOE for the HSRP targeting. Two-sample t-tests are then used to check if there are still significant differences in the means of observable characteristics for both groups (see section 5.6). Second, we conduct a DD estimation in which health outcomes are defined conditional on the propensity score.

While we use district-level social and economic indicators to estimate the propensity score, facility-level characteristics are used as additional covariates later in the DD estimations. For each of the outcomes, the results are reported for two study periods: 2000-2005 and 2005-2008.

5.4 DATA

We draw on data from the Egypt DHS to calculate facility-level health outcomes, information from MOH to capture interventions under the HSRP and construct a set of facility-level controls, and data from Egypt's Population and Housing Census to calculate a set of district-level social and economic controls. Whenever applicable, a number of regional dummies is included as well in our analyses.

Dependent Variables. We calculate the same health outcomes used in chapter 4 to estimate the effect of introducing user fees. These are the outcomes that reflect the utilization of family planning, ANC and delivery care services, women's access to health care, and the health status of children. As discussed in chapter 4, these are the outcomes that are expected to be affected by introducing user fees.

To calculate these outcomes, we follow the same steps discussed in chapter 3. First, we spatially join each woman interviewed in each of the DHS waves 2000, 2005 and 2008 to her nearest mapped health facility using Quantum GIS 2.8.2. Second, for each of the Egypt DHS waves, we use Stata 12.0 to recode and calculate health outcomes at the facility level. Third, we combine our outcomes of interest in a panel. A detailed description of the definitions and sources of the health outcomes included in our analyses is provided in chapter 4 (see Table 4.2).

Explanatory Variables. The main explanatory variable included in our analyses is $cont_{it}$. The variable draws on information from Egypt's MOH regarding whether and when a facility is non-contracted or contracted. Shifting from non-contracted to contracted, a facility introduces user fees accompanied by two quality improvement interventions: facility accreditation and PBF of health providers.

To eliminate any unobserved heterogeneity, we include a set of facility-level, district-level, and regional covariates as controls in our models whenever possible. These are the same facility-level characteristics, district-level social and economic characteristics, and regional dummies discussed in chapter 3 (see Table 3.4).

5.5 RESULTS

In this section, we report the descriptive statistics for the sample used as well as the estimated effects of introducing user fees and quality improvements under the HSRP on family planning, maternal health and child health outcomes. As previously noted,

for each health outcome, we report the results of estimating three DD specifications and a DD PSM model specification denoted by specification (4). This is our preferred specification in which all control variables are included either as matchers or as additional covariates (see Table 3.4). Our analysis of the effect is mainly based on the results of this specification.

5.5.1 Descriptive Statistics

Differences in the facility and the district characteristics between contracted and non-contracted health facilities are highlighted in Tables 5.1 and 5.2, respectively. Two-sample t-tests are used to check whether the means of the two groups differ significantly.

We observe significant differences between contracted and non-contracted facilities with respect to six out of 10 facility-level characteristics during the period 2000-2005. Contracted facilities employ more practitioners (*pract*), specialists (*spec*), pharmacists (*pharm*), lab technicians (*labtech*) and health observers (*obs*) compared to non-contracted facilities. The building condition (*inf*) of contracted facilities is also significantly better. Similar differences are observed between contracted and non-contracted facilities during the study period 2005-2008 (Table 5.1).

In parallel, contracted facilities are located in districts that significantly have lower unemployment rates (*unemp*) and bigger population size (*pop*) during the study period 2000-2005. During the period 2005-2008, we observe that contracted facilities are located in districts that significantly have lower income dependency (*incdep*), better accessibility to electricity (*elect*) and bigger population size (*pop*) (Table 5.2).

Table 5.2 provides evidence that the actual targeting of the HSRP does not strictly follow the socio-vulnerability index. The differences in the district characteristics captured between contracted and non-contracted facilities are inconsistent with the targeting criteria of the HSRP. These differences could bias our estimation results. Thus, we must match facilities before estimating our DD models as previously discussed.

Table 5.1: Two-sample t-test of facility characteristics of contracted and non-contracted facilities

	2000-2005			2005-2008		
	Non-contracted	Contracted	Difference	Non-contracted	Contracted	Difference
Practitioners	3.367	5.956	-2.589***	3.137	4.022	-0.885***
			(0.581)			(0.314)
Specialists	0.530	1.324	-0.793***	0.522	0.731	-0.209*
			(0.206)			(0.121)
Pharmacists	3.931	7.221	-3.290***	3.723	4.762	-1.039***
			(0.692)			(0.396)
Nurses	11.853	10.074	1.780	11.937	8.776	3.161***
			(1.600)			(0.946)
Lab technicians	1.290	2.103	-0.813***	1.336	1.296	0.040
			(0.205)			(0.122)
X-ray technicians	0.177	0.250	-0.073	0.193	0.126	0.067
			(0.103)			(0.061)
Health observers	1.309	2.235	-0.926***	1.330	1.426	-0.096
			(0.186)			(0.111)
Social workers	0.437	0.309	0.128	0.413	0.256	0.158
			(0.199)			(0.117)
Building condition	1.630	1.857	-0.228***	1.605	1.836	-0.231***
			(0.081)			(0.048)
Population coverage	3.305	2.554	0.751	3.269	2.592	0.677
			(1.364)			(0.820)

Standard errors are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

Table 5.2: Two-sample t-test of district characteristics of contracted and non-contracted facilities

	2000-2005			2005-2008		
	Non-contracted	Contracted	Difference	Non-contracted	Contracted	Difference
Illiteracy	32.376	31.445	0.931	32.672	32.078	0.594
			(1.268)			(0.719)
Unemployment	9.703	8.514	1.189**	9.635	9.685	-0.051
			(0.547)			(0.317)
Income dependency	5.228	3.185	2.043	5.602	2.524	3.078***
			(1.911)			(1.102)
Inaccessibility to electricity	1.285	0.842	0.443	1.357	0.916	0.441*
			(0.422)			(0.249)
Inaccessibility to potable water	4.390	3.845	0.545	4.528	4.496	0.033
			(0.856)			(0.499)
Family size	4.330	4.309	0.021	4.334	4.397	-0.063**
			(0.046)			(0.026)
HH overcrowding	1.148	1.178	-0.030**	1.141	1.207	-0.067***
			(0.013)			(0.007)
Population size	31.646	39.952	-8.306***	31.072	35.823	-4.751***
			(2.022)			(1.135)

Standard errors are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

Tables 5.3 and 5.4 provide the descriptive statistics of the health outcomes of facilities before and after contracting over two time periods: 2000-2005 and 2005-2008. These outcomes are calculated based on the answers of women interviewed in the Egypt DHS after identifying who are located in catchment areas of contracted facilities (treatment group) and who are located in catchment areas of non-contracted facilities (control group). We observe that treatment and control groups of facilities are matched during both study periods 2000-2005 and 2005-2008. We do not observe statistically significant differences at the baseline between contracted and non-contracted facilities during both periods.

5.5.2 Estimated Effects of Introducing User Fees and Quality Improvements

During the study period 2000-2005, introducing user fees and quality improvements had positive effects on ANC and women's access to health care; but no effects on family planning, delivery care and child health status. No negative effects are observed with respect to any of our health outcomes of interest (Table 5.5).

Having access to a contracted facility was associated with a higher likelihood of receiving ANC by skilled health personnel (*ancprov*), receiving four or more ANC visits (*anc4*) and receiving iron supplements during pregnancy (*anciron*) during the period 2000-2005. Table 5.5 shows that *ancprov*, *anc4* and *anciron* increased by 12 ppts, 7 ppts and 10 ppts, respectively, between 2000 and 2005, among women with access to contracted facilities compared to those with access to non-contracted facilities. A possible explanation of the multiple positive effects on ANC is that the ANC dimension was subject to substantial quality improvements under the facility accreditation program.

In addition, we observe that introducing user fees and quality improvements had a significant positive effect on women's access to health care. Having access to a contracted facility is associated with a lower likelihood of reporting money as an impediment to access care. Table 5.5 shows that the proportion of women with access to contracted facilities, who report getting money for treatment (*accmon*) as a problem in accessing health care decreased significantly by 7 ppts between 2000 and 2005. The positive effect captured with respect to *accmon* suggests that access to health care is enhanced, not impeded, when quality improvements accompany user fee introduction.

Table 5.3: Descriptive statistics of health outcomes, 2000-2005

		Baseline (2000)			Follow up (2005)		
		Non-contracted	Contracted	Difference	Non-contracted	Contracted	Difference
Family planning	Modern contraceptive prevalence	48.269	48.858	0.589	50.381	50.546	0.166
		(0.805)	(3.616)	(3.557)	(0.655)	(3.051)	(3.037)
ANC	ANC by skilled health personnel	54.083	52.688	-1.395	68.370	77.422	9.052*
		(1.104)	(6.074)	(6.466)	(0.813)	(5.051)	(5.152)
	4+ visits	38.024	38.323	0.299	57.441	60.932	3.491
		(1.083)	(7.262)	(7.176)	(0.843)	(3.534)	(3.818)
	Iron supplementation	25.737	28.211	2.474	48.713	60.305	11.592***
		(0.761)	(6.286)	(6.260)	(0.893)	(4.149)	(4.254)
Delivery care	Institutional delivery	49.228	55.158	5.930	62.623	68.266	5.643
		(1.277)	(8.073)	(8.139)	(1.180)	(4.586)	(4.806)
	Skilled-assisted delivery	61.768	62.474	0.706	73.456	79.546	6.090
		(1.360)	(7.162)	(7.219)	(0.921)	(4.004)	(4.261)
	C-section delivery	10.706	13.170	2.464	19.348	24.233	4.885
		(0.499)	(3.097)	(3.104)	(0.644)	(3.653)	(3.716)
Access to care	Money barrier reported	28.581	32.219	3.638	30.940	27.049	-3.891
		(0.794)	(3.404)	(3.326)	(0.740)	(3.103)	(3.209)
Child health status	Under-5 mortality	4.260	4.504	0.244	3.422	2.918	-0.503
		(0.255)	(1.317)	(1.348)	(0.180)	(1.194)	(1.216)

Bootstrapped standard errors are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

Table 5.4: Descriptive statistics of health outcomes, 2005-2008

		Baseline (2005)			Follow up (2008)		
		Non-contracted	Contracted	Difference	Non-contracted	Contracted	Difference
Family planning	Modern contraceptive prevalence	49.989	48.180	-1.809	51.192	50.877	-0.315
		(0.519)	(1.859)	(2.017)	(0.715)	(1.970)	(2.018)
ANC	ANC by skilled health personnel	68.185	66.509	-1.676	73.161	62.842	-10.319***
		(1.086)	(2.530)	(2.812)	(0.897)	(3.071)	(3.329)
	4+ visits	56.966	55.112	-1.854	64.282	56.422	-7.860**
		(0.945)	(2.403)	(2.388)	(1.109)	(3.257)	(3.568)
	Iron supplementation	48.130	50.835	2.705	38.063	45.404	7.341***
		(0.961)	(2.380)	(2.596)	(0.978)	(2.247)	(2.307)
Delivery care	Institutional delivery	62.120	59.766	-2.355	70.111	68.683	-1.428
		(1.258)	(2.871)	(3.052)	(0.928)	(3.357)	(3.551)
	Skilled-assisted delivery	72.893	74.769	1.876	78.311	77.246	-1.065
		(0.987)	(1.482)	(1.895)	(1.048)	(3.187)	(3.363)
	C-section delivery	18.879	18.513	-0.366	27.790	25.430	-2.360
		(0.835)	(1.625)	(1.733)	(0.689)	(2.077)	(2.273)
Access to care	Money barrier reported	30.298	30.327	0.029	47.056	51.153	4.096
		(0.617)	(1.352)	(1.578)	(0.844)	(3.163)	(3.305)
Child health status	Under-5 mortality	3.518	3.614	0.096	2.241	2.794	0.553
		(0.246)	(0.548)	(0.621)	(0.223)	(0.542)	(0.561)

Bootstrapped standard errors are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

Table 5.5: Estimated effects of introducing user fees and quality improvements, 2000-2005

	Outcome	DD			DD PSM
		(1)	(2)	(3)	(4)
Family planning	Modern contraceptive prevalence	-0.423	-0.210	0.386	1.761
		(4.887)	(4.165)	(4.266)	(2.410)
ANC	ANC by skilled health personnel	10.447	5.164	7.133	11.830***
		(8.193)	(5.631)	(5.111)	(4.301)
	4+ visits	3.193	-0.618	0.977	7.086*
		(7.976)	(6.784)	(6.335)	(3.898)
	Iron supplementation	9.118	8.767	8.795	9.522**
		(7.927)	(6.358)	(6.246)	(3.833)
Delivery care	Institutional delivery	-0.287	-4.013	-1.172	-1.930
		(8.587)	(8.382)	(7.413)	(4.672)
	Skilled-assisted delivery	5.384	-0.397	2.042	2.940
		(7.901)	(8.715)	(7.530)	(4.405)
	C-section delivery	2.421	0.973	1.287	2.595
		(4.810)	(4.075)	(4.769)	(2.670)
Access to care	Money barrier reported	-7.529*	-2.349	-2.524	-7.034**
		(4.491)	(5.379)	(4.935)	(2.932)
Child health status	Under-5 mortality	-0.747	-0.527	-0.658	-1.142
		(1.905)	(1.353)	(1.388)	(0.940)

Each row represents a separate regression. Standard errors are reported in parentheses. Bootstrapped standard errors are reported for specifications (1) to (3). *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

We do not observe any significant effects during the period 2000-2005 on outcomes of family planning, delivery care and child health status (Table 5.5).

However, Table 5.6 indicates that the positive effects observed during the study period 2000-2005 were reversed during the period 2005-2008. This change in the effects is justified by the fact that the phased introduction of user fees started in the first quarter of 2004. Thus, our second study period is more likely to reflect the effects of user fees while the first period is more likely to reflect the effects of quality improvements through facility accreditation and PBF of health providers. The results of the period 2005-2008 provide evidence that user fees drove down utilization in contracted facilities and probably shifted utilization towards non-contracted facilities. Although the poor are supposed to be officially exempt from any fees at the point of service, this exemption is not essentially known to the population (World Bank, 2010). According to a facility survey carried out by the World Bank as a census of all 362 public PHC facilities in Alexandria and Menoufia and a HH survey of a total of 5,471 HHs in their catchment area, 97 percent of the respondents have never heard of the payment exemption for the poor in contracted facilities.

Table 5.6: Estimated effects of introducing user fees and quality improvements, 2005-2008

	Outcome	DD			DD PSM
		(1)	(2)	(3)	(4)
Family planning	Modern contraceptive prevalence	1.493	0.349	-0.445	2.279
		(2.961)	(2.323)	(2.204)	(2.428)
ANC	ANC by skilled health personnel	-8.643**	-11.17***	-12.54***	-7.551*
		(4.068)	(3.626)	(3.898)	(4.128)
	4+ visits	-6.006	-7.572	-9.004**	-4.031
		(4.112)	(4.632)	(4.411)	(4.151)
	Iron supplementation	4.636	3.408	3.474	4.641
		(3.385)	(3.790)	(3.801)	(3.941)
Delivery care	Institutional delivery	0.927	-2.237	-3.653	-2.178
		(4.888)	(4.187)	(4.045)	(4.190)
	Skilled-assisted delivery	-2.941	-6.107*	-7.161**	-0.073
		(3.919)	(3.408)	(3.348)	(3.961)
	C-section delivery	-1.994	-2.871	-3.790	-3.068
		(2.821)	(3.233)	(3.218)	(2.963)
Access to care	Money barrier reported	4.067	3.483	3.983	9.353***
		(2.923)	(3.669)	(3.651)	(3.022)
Child health status	Under-5 mortality	0.457	0.491	0.553	0.952
		(0.802)	(1.017)	(0.993)	(1.004)

Each row represents a separate regression. Standard errors are reported in parentheses. Bootstrapped standard errors are reported for specifications (1) to (3). *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

The change in the effects could also be justified by the significant reduction of incentive payments to health providers due to funding issues of the FHFs. World Bank (2010) reports that the average utilization rate per person per year decreased, from 25 daily encounters per doctor in 2006 to 12 daily encounters in 2009, possibly because of decreased funding for incentives. This finding provides evidence on the effect of provider incentives on service utilization and patient satisfaction.

5.6 ROBUSTNESS CHECKS

In this section, we run several checks to verify the robustness of our main results reported in the previous section. First, a placebo test is carried out to verify the functional form of the DD set-up. Second, we assess the quality of matching facilities using two-sample t-tests. Third, we conduct a sensitivity analysis for our DD PSM estimates.

5.6.1 Placebo Test

We follow the placebo test of Bertrand *et al.* (2004) as in chapter 3 to ensure that the functional form of the DD set-up is properly specified. If so, no effects on the health outcomes of contracted facilities should be observed before contracting as we define a "false" lagged contracting intervention. The data of the study period 2000-2005 is used to verify the results of the period 2005-2008, where facilities that are contracted after 2005 are defined as treated and facilities that are not contracted after 2005 are defined as control for the period 2000-2005. We could not run the placebo test using our preferred model's specification (4) due to data limitation. Alternatively, we use the model's specification (3).

The results of our placebo test reported in Table 5.7 indicate that the treatment estimates are not significantly different from zero for all health outcomes. That is, differences between contracted and non-contracted facilities reported in our main results only emerged after contracting.

Table 5.7: Estimated effects of placebo contracting, 2000-2005

	Outcome	Difference (baseline)	Difference (follow-up)	DD (3)
Family planning	Modern contraceptive prevalence	-7.269*	3.210	10.479
				(6.397)
ANC	ANC by skilled health personnel	-19.073	-10.510	8.563
				(53.932)
	4+ visits	-13.851	-14.609	-0.758
				(27.636)
	Iron supplementation	-21.195	-13.857	7.338
				(20.424)
Delivery care	Institutional delivery	-16.768	-33.036	-16.268
				(13.643)
	Skilled-assisted delivery	-26.650	-31.684	-5.034
				(42.505)
	C-section delivery	-29.138	-26.366	2.772
				(8.407)
Access to care	Money barrier reported	0.482	8.686	8.204
				(10.992)
Child health status	Under-5 mortality	2.334	3.113	0.778
				(5.280)

Each row represents a separate regression. The covariates are the facility characteristics, district socio-economic indicators and regional dummies from Table 3.4. Standard errors are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

5.6.2 Quality of Matching

We use the balancing two-sample t-test of the difference in means of covariates across matched samples of facilities to check whether observable characteristics are balanced in the matched sample.

The results of the t-test reported in Table 5.8 indicate that matching on the propensity score is successful as there are no systematic differences in general at the baseline in the means of observed characteristics between contracted and non-contracted health facilities.

5.6.3 Sensitivity of Results

We follow the same steps discussed in chapter 3 to inspect the sensitivity of our main estimation results reported in section 5.5 to the type of the Kernel function, the bandwidth of the Kernel function and the estimation method of the propensity score (see subsection 3.6.3).

The results of our sensitivity checks reported in Tables 5.9 and 5.10 provide evidence that our main results are not sensitive in general to alternative types of the Kernel function. More importantly, our main results are not sensitive to alternative bandwidths of the Kernel function (Tables 5.11 and 5.12). Moreover, the results of our checks reported in Table 5.13 indicate that our main results are not sensitive to alternative estimation methods of the propensity score.

Table 5.8: Balancing two-sample t-test, 2005-2008

	Difference in mean covariates between contracted and non-contracted facilities							
	Illiteracy	Un-employment	Income dependency	In-accessibility to electricity	In-accessibility to potable water	Family size	HH overcrowding	Population size
Family planning								
Modern contraceptive prevalence	0.508	0.134	-0.510	-0.020	0.063	0.011	-0.002	4.040***
ANC								
ANC by skilled health personnel	-1.180	0.416	-1.209	-0.051	-0.356	-0.002	-0.013	4.908***
4+ visits	-1.180	0.416	-1.209	-0.051	-0.356	-0.002	-0.013	4.908***
Iron supplementation	-1.180	0.416	-1.209	-0.051	-0.356	-0.002	-0.013	4.908***
Delivery care								
Institutional delivery	0.273	0.204	-1.399	-0.065	-0.191	-0.004	-0.009	5.761***
Skilled-assisted delivery	-1.139	0.431	-1.284	-0.051	-0.350	-0.001	-0.013	4.958***
C-section delivery	-1.175	0.415	-1.199	-0.051	-0.357	-0.001	-0.013	4.903***
Access to care								
Money barrier reported	0.508	0.134	-0.510	-0.020	0.063	0.011	-0.002	4.040***
Child health status								
Under-5 mortality	0.324	0.175	-1.137	-0.047	0.035	0.008	-0.008	5.201***

Means and t-test are estimated by linear regression. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

Table 5.9: Sensitivity to the type of the Kernel function, 2000-2005

	Outcome	Main results	Type of function			
			Gaussian	Biweight	Uniform	Tricube
Family planning	Modern contraceptive prevalence	1.761 (2.410)	1.496 (2.452)	1.868 (2.412)	1.515 (2.408)	1.653 (2.409)
ANC	ANC by skilled health personnel	11.830*** (4.301)	10.654** (4.387)	11.999*** (4.304)	11.550*** (4.293)	11.572*** (4.294)
	4+ visits	7.086* (3.898)	6.768* (3.979)	7.232* (3.901)	6.872* (3.892)	6.870* (3.892)
	Iron supplementation	9.522** (3.833)	9.130** (3.910)	9.619** (3.840)	9.281** (3.820)	9.319** (3.822)
Delivery care	Institutional delivery	-1.930 (4.672)	-2.362 (4.803)	-1.737 (4.676)	-2.089 (4.664)	-2.105 (4.664)
	Skilled-assisted delivery	2.940 (4.405)	2.230 (4.534)	3.031 (4.406)	2.951 (4.402)	2.900 (4.401)
	C-section delivery	2.595 (2.670)	2.305 (2.744)	2.488 (2.677)	2.707 (2.664)	2.712 (2.663)
Access to care	Money barrier reported	-7.034** (2.932)	-5.869** (2.977)	-7.162** (2.937)	-6.889** (2.921)	-6.937** (2.932)
Child health status	Under-5 mortality	-1.142 (0.940)	-0.963 (0.956)	-1.188 (0.938)	-1.111 (0.941)	-1.106 (0.941)

Each row represents a separate regression. Standard errors are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

Table 5.10: Sensitivity to the type of the Kernel function, 2005-2008

	Outcome	Main results	Type of function			
			Gaussian	Biweight	Uniform	Tricube
Family planning	Modern contraceptive prevalence	2.279 (2.428)	2.167 (2.486)	1.886 (2.410)	2.965 (2.477)	2.827 (2.459)
ANC	ANC by skilled health personnel	-7.551* (4.128)	-7.593* (4.028)	-7.548* (4.136)	-7.851* (4.163)	-7.652* (4.132)
	4+ visits	-4.031 (4.151)	-4.475 (4.055)	-3.964 (4.160)	-4.514 (4.175)	-4.171 (4.153)
	Iron supplementation	4.641 (3.941)	3.816 (3.805)	4.630 (3.962)	4.684 (3.937)	4.613 (3.924)
Delivery care	Institutional delivery	-2.178 (4.190)	-1.727 (4.335)	-2.269 (4.180)	-2.133 (4.190)	-2.101 (4.197)
	Skilled-assisted delivery	-0.073 (3.961)	-2.129 (3.884)	-0.168 (3.971)	-0.303 (3.975)	0.004 (3.957)
	C-section delivery	-3.068 (2.963)	-4.141 (2.944)	-3.210 (2.967)	-2.197 (2.993)	-2.751 (2.965)
Access to care	Money barrier reported	9.353*** (3.022)	9.009*** (3.088)	8.729*** (3.020)	10.063*** (3.033)	10.044*** (3.027)
Child health status	Under-5 mortality	0.952 (1.004)	0.621 (1.023)	1.012 (1.004)	0.811 (1.002)	0.862 (1.004)

Each row represents a separate regression. Standard errors are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

Table 5.11: Sensitivity to the bandwidth of the Kernel function, 2000-2005

	Outcome	Main results	Bandwidth	
			0.05	0.1
Family planning	Modern contraceptive prevalence	1.761	1.899	1.556
		(2.410)	(2.409)	(2.427)
ANC	ANC by skilled health personnel	11.830***	11.974***	11.531***
		(4.301)	(4.304)	(4.311)
	4+ visits	7.086*	7.201*	6.944*
		(3.898)	(3.900)	(3.907)
	Iron supplementation	9.522**	9.629**	9.383**
		(3.833)	(3.838)	(3.837)
Delivery care	Institutional delivery	-1.930	-1.815	-1.986
		(4.672)	(4.675)	(4.684)
	Skilled-assisted delivery	2.940	2.985	3.080
		(4.405)	(4.405)	(4.422)
	C-section delivery	2.595	2.529	2.633
		(2.670)	(2.674)	(2.677)
Access to care	Money barrier reported	-7.034**	-7.141**	-5.912**
		(2.932)	(2.935)	(2.946)
Child health status	Under-5 mortality	-1.142	-1.181	-1.089
		(0.940)	(0.937)	(0.942)

Each row represents a separate regression. Standard errors are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

Table 5.12: Sensitivity to the bandwidth of the Kernel function, 2005-2008

	Outcome	Main results	Bandwidth	
			0.05	0.1
Family planning	Modern contraceptive prevalence	2.279	1.917	2.562
		(2.428)	(2.405)	(2.469)
ANC	ANC by skilled health personnel	-7.551*	-7.473*	-7.662*
		(4.128)	(4.127)	(4.047)
	4+ visits	-4.031	-3.874	-4.487
		(4.151)	(4.151)	(4.073)
	Iron supplementation	4.641	4.663	4.229
		(3.941)	(3.959)	(3.821)
Delivery care	Institutional delivery	-2.178	-2.275	-1.545
		(4.190)	(4.179)	(4.294)
	Skilled-assisted delivery	-0.073	-0.102	-1.663
		(3.961)	(3.965)	(3.888)
	C-section delivery	-3.068	-3.331	-3.305
		(2.963)	(2.961)	(2.942)
Access to care	Money barrier reported	9.353***	8.810***	9.191***
		(3.022)	(3.018)	(3.048)
Child health status	Under-5 mortality	0.952	1.005	0.697
		(1.004)	(1.004)	(1.013)

Each row represents a separate regression. Standard errors are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

Table 5.13: Sensitivity to the estimation method of the propensity score

	Outcome	2000-2005		2005-2008	
		Probit	Logit	Probit	Logit
Family planning	Modern contraceptive prevalence	1.761 (2.410)	1.757 (2.414)	2.279 (2.428)	1.066 (2.504)
ANC	ANC by skilled health personnel	11.830*** (4.301)	11.781*** (4.301)	-7.551* (4.128)	-7.985* (4.114)
	4+ visits	7.086* (3.898)	7.041* (3.898)	-4.031 (4.151)	-4.101 (4.143)
	Iron supplementation	9.522** (3.833)	9.486** (3.831)	4.641 (3.941)	4.120 (3.961)
Delivery care	Institutional delivery	-1.930 (4.672)	-1.986 (4.671)	-2.178 (4.190)	0.767 (4.357)
	Skilled-assisted delivery	2.940 (4.405)	2.942 (4.405)	-0.073 (3.961)	-0.262 (3.942)
	C-section delivery	2.595 (2.670)	2.598 (2.669)	-3.068 (2.963)	-2.939 (2.950)
Access to care	Money barrier reported	-7.034** (2.932)	-6.976** (2.938)	9.353*** (3.022)	9.616*** (3.152)
Child health status	Under-5 mortality	-1.142 (0.940)	-1.127 (0.941)	0.952 (1.004)	0.701 (1.039)

Each row represents a separate regression. Standard errors are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

5.7 CONCLUSION

In this chapter, we collapse data from the Egypt DHS waves of 2000, 2005 and 2008 at the facility level to complement our results of introducing user fees obtained in chapter 4. Our aim is to investigate how combining user fee introduction and quality improvements affects the utilization of family planning, ANC and delivery care services, women's access to health care, and child health status. DD and DD Kernel PSM are used in this context.

If accompanied by quality improvements, introducing user fees is associated with positive effects on the utilization of ANC services and on women's access to health care during the study period 2000-2005. Having access to a contracted facility was associated with a higher likelihood of receiving ANC by skilled health personnel, receiving four or more ANC visits and receiving iron supplements during pregnancy. These positive effects on ANC probably reflect substantial quality improvements in the ANC dimension under the facility accreditation program. Moreover, the phased introduction of user fees started in the first quarter of 2004. Thus, we do not expect this study period to fully capture the effects of user fee introduction. In parallel, we find that having access to a contracted facility is associated with a lower likelihood of

reporting money as an impediment to access care during the period 2000-2005. No effects are observed during this period with respect to family planning, delivery care and child health status.

The positive effects observed during the study period 2000-2005 were reversed during the period 2005-2008. This change in effects was probably driven by the introduction of user fees in 2004 together with the low level of public awareness of the exemption policy of the poor. Moreover, the significant reduction of incentive payments to health providers due to funding issues of the FHFs could have affected the quality of care.

The findings of this chapter provide evidence that, even when accompanied by quality improvements, introducing user fees in low- and middle-income settings can have negative effects on access to and utilization of health care. These negative effects are expected to be more significant in light of unpublicized exemptions and unsustainable quality improvements.

6. THE EFFECT OF DISCONTINUING PROVIDER INCENTIVES ON FAMILY PLANNING AND MATERNAL AND CHILD HEALTH

6.1 INTRODUCTION

In 2001, a performance-based financing (PBF) scheme -also known as pay-for-performance (P4P)- was integrated in the Health Sector Reform Program (HSRP) of Egypt. According to this scheme, the Family Health Fund (FHF) is entitled to pay monthly incentives to healthcare providers in contracted facilities based on pre-specified performance criteria. To qualify for financial incentives, health facilities are required to underscore pre-determined standards of 11 indicators that cover several aspects of service provision, whether curative or preventive, and maintains efficiency and quality. At the end of 2008, however, financial incentives to providers in contracted facilities were discontinued. The PBF scheme was replaced by the conventional "fee-for-service" mechanism to ensure the financial sustainability of FHFs.

Provider incentives are increasingly advocated as effective means to change the behavior of providers, and, consequently, improve health outcomes. While there is considerable enthusiasm for the intervention, there is little rigorous evidence on its effectiveness in low- and middle-income settings, particularly when implemented at scale (Das *et al.*, 2016; Witter *et al.*, 2012; Eldridge & Palmer, 2009). The existing evidence is more limited on the effects of discontinuing rather than introducing incentives, even in high-income settings. Moreover, the available studies focus on process indicators, quality, etc., rather than patient outcomes. No evidence is found as well on the effects of discontinuing provider incentives in any low- or middle-income country. As many PBF schemes have been operating for several years, the need to improve policy design requires an insight into the effects of discontinuing these incentives either partly or totally. Such discontinuation may be stimulated by changes in policy priorities, by the ineffectiveness of some schemes, or by the unintended dysfunctional consequences of other schemes.

In this chapter, we fill the gap in the literature by investigating the effects of discontinuing provider incentives in Egypt after being in operation for more than five years. We estimate the effects of replacing PBF by "fee-for-service" on family planning, maternal health and child health outcomes that reflect the health services targeted by the PBF scheme as well as the quality of these services in contracted facilities. Generalized least squares (GLS) random-effects (RE) and fixed-effects (FE) are used in this context.

The chapter proceeds as follows. Section 6.2 provides background information for our analysis; section 6.3 discusses the econometric method used as well as considerations for model choice; section 6.4 constructs our dependent and explanatory variables; section 6.5 reviews the descriptive statistics and presents the main results of our estimations; section 6.6 reports the results of some robustness checks; and section 6.7 concludes.

6.2 BACKGROUND

In this section, we give an overview of the design of the PBF scheme, discuss the anticipated effect of discontinuing provider incentives and review evidence on the effect of introducing or discontinuing provider incentives in low- and middle-income countries.

6.2.1 PBF Scheme Design

To increase the quality of, as well as access to and use of health services, a PBF scheme was integrated in the HSRP in 2001 to complement the facility accreditation program. According to this scheme, FHFs paid financial incentives on a monthly basis to healthcare providers in contracted health facilities, who deliver the BBP, based on pre-defined performance criteria. All workers in contracted facilities were eligible to receive incentives, including doctors, nurses, technicians, administrators, other health workers and support staff. These workers received a base salary, which was typically low, in addition to an incentive payment that could reach up to 275 percent of their base salary.

To qualify for financial incentives, facilities were required to underscore pre-determined standards of 11 indicators that were selected by MOH to reflect various aspects of service provision. Through these indicators, the PBF scheme sought to

address priority health concerns in Egypt, including maternal and child health, reproductive health/family planning, tuberculosis (TB), and immunization. A list of the 11 indicators is provided in Table 6.1 together with the required target assigned to each indicator.

Table 6.1: Performance Indicators of the PBF Scheme

	Indicator	Target
1	Number of visits per day per physician	10-24
2	Number of drugs per visit	Less than 2
3	Rate of patient referral to the district hospital	1-8%
4	Rate of completion of visit encounter forms	More than 98%
5	Patient satisfaction rate	More than 90%
6	Rate of completion of medical records data	More than 90%
7	Years of protection*	More than 50%
8	Number of children fully vaccinated in the catchment area	95%
9	Patient waiting time	Less than 20 minutes
10	Number of ANC visits per pregnant woman	More than half a visit per month
11	Adherence to medical protocols	More than 98%

*Years of protection is the estimated protection provided by contraceptive methods during a one-year period, based upon the volume of all contraceptives sold or distributed free of charge to beneficiaries during that period.
Source: Egypt's MOH

The higher the performance of a facility with respect to these indicators, the higher the incentives that were paid by the FHF to this facility. According to the PBF scheme, if facilities achieved more than 75 percent of the indicator targets, they would receive 100 percent of the financial incentives; if facilities achieved 61-75 percent, they would receive 50 percent of the incentives. Facilities that achieved less than 61 percent would receive no incentives.

When a facility met certain targets, the relevant FHF would make a cash payment to the facility manager, who then distributed the incentives to the staff involved in achieving the target. To determine which staff participated in achieving the target, each facility had its own pre-determined protocol that was based on a point system. This point system was, in turn, based on number of variables, such as qualifications, experience, number of days worked and efforts made to achieve the indicators in each area. The total cash payments made to a facility was divided by the sum of the points earned by the staff and multiplied by the number of points for each worker. This determined the amount of cash payment each worker received each month.

The PBF scheme was replaced by the conventional "fee-for-service" mechanism by the end of 2008 to ensure the financial sustainability of the FHF.

6.2.2 Anticipated Effect of Discontinuing Provider Incentives

Egypt's MOH introduced a PBF scheme to ensure that the provision of the BBP is based upon competition between different healthcare providers. Under this scheme, performance-based supply-side financial incentives were disbursed to providers in contracted facilities to strengthen their accountability to performance targets and induce desirable health outcomes. Economic theory suggests that PBF could alter the supply of health services. The contract theory underlying PBF schemes is traced to the principal-agent literature. According to the conventional principal-agent model, the PBF scheme is an intervention to stimulate efficient provider behavior through financial incentives. The ultimate goal is to improve the quality and efficiency of health care. We start this subsection by reviewing a simple model of healthcare provision under an incentive payment scheme to understand how providers are expected to behave if paid based on their performance. Accordingly, we discuss the anticipated effect of discontinuing financial incentives to providers.

A Model of Healthcare Provision under an Incentive Payment Scheme. Contract theory suggests that contracts could be incentivized to induce agents –providers in the context of this study– to perform and/or not perform actions. Incentive-based contracts are used to address information asymmetries in both provider-patient and payer-provider relationships. A perfect agent should perform actions that maximize the wellbeing of the patients. However, since a typical provider works to maximize his/her own utility, contracts should be designed to promote targeted health outcomes.

We attempt to formalize this idea in a simplified principal-agent framework, where a healthcare provider (the agent) could use his/her "effort" to improve the quality of health services provided. However, since this "effort" is unobservable to the purchaser of services or the payer (the principal such as Egypt's MOH), the payer should reimburse the provider according to contracted outcomes, not according to his/her choice of efforts.

We consider a healthcare service contract problem in which a purchaser of services (Egypt's MOH) offers a contract to a healthcare provider (PHC facility) to deliver outpatient services. We, first, assume that two health outcomes, y^c and y^{nc}, could be

generated by a provider. y^c denotes a contractible outcome against which a provider is being compensated and y^{nc} denotes a non-contractible outcome.

We further assume that a provider chooses two actions, exerting two types of efforts, e_1 and e_2, at cost $c(e_1, e_2)$. Both efforts are costly to exert and both are unobservable except by the provider. A provider could produce each of the two health outcomes as follows:

$$y^c = f(e_1, e_2) + \varepsilon \tag{6.1.a}$$

$$y^{nc} = g(e_1, e_2) + \eta \tag{6.1.b}$$

We assume that the provider has a separable utility function u that is increasing in compensation w and decreasing in efforts e_1 and e_2.

$$u(w, e_1, e_2) = w - c(e_1, e_2) \tag{6.2}$$

where u is concave for a risk-averse provider, convex for a risk-seeking provider and linear for a risk-neutral provider.

To formalize the idea that providers are paid based on their performance with respect to contractible health outcomes, we assume that the compensation of a provider w is a linear function of the amount of the contractible health outcome y^c produced, which is in turn a function of efforts e_1 and e_2:

$$w = \alpha + \beta y^c \tag{6.3}$$

Thus, the provider chooses his/her efforts to maximize utility less the cost of effort. Substituting w in equation (6.2) by equation (6.3), we obtain the initial provider's optimization program given by:

$$e^* = arg\,max_e\ \alpha + \beta y^c - c(e_1, e_2) \tag{6.4}$$

Substituting y^c in equation (6.4) by equation (6.1.a), we obtain the final provider's optimization program given by:

$$e^* = arg\ max_e\ \alpha + \beta[f(e_1, e_2) + \varepsilon] - c(e_1, e_2) \tag{6.5}$$

For a given w, the first-order conditions for the maximization of equation (6.5) with respect to the amounts of e_1 and e_2 are:

$$\frac{dc}{de_1} = \beta \frac{df}{de_1} \text{ and } \frac{dc}{de_2} = \beta \frac{df}{de_2} \tag{6.6}$$

The optimal level of effort is chosen by the provider to equalize the marginal cost of effort and the marginal increase in the amount of the contractible health outcome y^c produced as a result of this effort. We expect the amount of the contractible health outcome y^c produced to increase as β increases. The actual level of effort a provider is willing to provide depends on the contract he/she has. In this regard, it is important to note that it is not feasible to predict the amount of the non-contractible health outcome y^{nc} produced as the function $g(e_1, e_2)$ does not appear in the first-order conditions.

During the process of designing incentive-based contracts for healthcare providers (agents), it is important for the purchaser of services (the principal) to seek to maximize its utility subject to the individual rationality and incentive compatibility constraints of the provider (the agent). This is a typical moral hazard situation.

Anticipated Effect. As previously discussed, contract theory suggests that financial incentives by the FHF would prompt higher quality performance of healthcare providers. Accordingly, we expect discontinuing the incentives to deter providers from using their "effort" to improve the quality of health services provided. This

expectation becomes even more justifiable given that these providers were offered these incentives before. Although it is beyond the scope of this study to observe the efforts of providers, we attempt to observe health outcomes and assess to what extent changes in observed outcomes are attributable to discontinuing provider incentives.

We expect discontinuing the incentives to have direct negative effects on the health services that were targeted by the PBF scheme (see Table 6.1). However, we also expect to observe an indirect negative effect on the quality and the utilization of PHC services in general in contracted facilities where incentives were discontinued.

6.2.3 Evidence on the Effect of Discontinuing/Introducing Provider Incentives

In this subsection, we review the available evidence on the effect of discontinuing provider incentives in high-income countries and introducing provider incentives in low- and middle-income countries.

6.2.3.1 Discontinuing Provider Incentives

Several databases were systematically searched on May 27, 2017 using EBSCOhost. These databases are: Academic Search Complete, Business Source Complete, CINAHL, EconLit, E-Journals, Health Policy Reference Center, MEDLINE, PsycINFO and SocINDEX. We used a combination of keywords to retrieve studies on the effect of discontinuing provider incentives: “performance based financing” or “performance based payment” or “pay for performance” or “provider incentive” or “incentives for provider” or “performance incentive”, together with “discontinu*” or “remov*” or “stop” or “ceas*” or “suspen*” or “terminat*”. This combination was used to search the abstracts of studies. This search yielded 88 studies after removing exact duplicates from the results. As the number of studies retrieved was initially small, we did not limit the results by excluding non-English studies, studies published before a particular year or studies conducted in high-income settings. The retrieved studies were screened based on title and abstract and the reference lists of the studies were also searched. A total of six studies were finally selected to be reviewed (see Table 6.2.a). All studies retrieved were conducted in high-income countries, four out of which in the United States, one in the United Kingdom and one in Italy.

Table 6.2.a: Evidence on the effectiveness of discontinuing provider incentives

Study	Intervention	Outcome measure	Reported effect on outcome
Benzer *et al.* (2013)	-Discontinuing performance-based incentives for selected inpatient quality measures	**Quality of care measures:**	
		Acute Coronary Syndrome	
		-Cardiology Involvement	None
		-Troponin Returned	None
		-Diagnostic Catheterization	None
		Heart Failure	
		-ACE-I or ARB	None
		-Weight Monitoring	None
		Pneumonia	
		-Timely Antibiotic	Positive
		-Pneumococcal Immunization	Positive
Boland *et al.* (2010)	-Discontinuing a radiologist PBF program	**Quality of care measures:**	
		Expedited finalized radiologist report turnaround times (RTAT)	
		-Examination completion (C) to final signature (F)	Positive
		-C to preliminary signature (P)	Positive
		-P to F	Positive
Fiorentini *et al.* (2013)	-Discontinuing financial incentives in PHC in hospitals	**Healthcare expenditure:**	
		-Avoidable hospital expenditure	None
		-Total hospital expenditure	None
Hysong *et al.* (2011)	-Discontinuing passive monitoring versus active assessment of clinical performance	**Quality of care measures:**	None or positive
		5 clinical areas common to ambulatory care:	
		-Screening	
		-Immunization	
		-Chronic care after acute myocardial infarction	
		-Diabetes mellitus	
		-Hypertension	

Study	Intervention	Outcome measure	Reported effect on outcome
Kontopantelis *et al*. (2014)	-Discontinuing financial incentives for aspects of care for patients with asthma, coronary heart disease, diabetes, stroke and psychosis	**Quality of care measures:** Influenza immunization -Asthma Lithium treatment monitoring -Psychosis Blood pressure monitoring -Coronary heart disease -Diabetes -Stroke Cholesterol concentration monitoring -Coronary heart disease -Diabetes Blood glucose monitoring -Diabetes	None
Lester *et al.* (2010)	-Discontinuing incentives for screening for diabetic retinopathy and for cervical cancer	**Quality of care measures:** -Screening for diabetic retinopathy -Screening for cervical cancer	 Negative Negative

ACE-I: Angiotensin-converting enzyme inhibitor. ARB: Angiotensin receptor blocker.

Five of the six studies reviewed investigated the effect of discontinuing incentives on recorded quality of care measures. Benzer *et al.* (2013) found that performance improvements that occurred in Veterans Health Administration (VA) medical centers in the United States for three common conditions (acute coronary syndrome, heart failure and pneumonia) were sustained for up to three years after performance-based incentives were discontinued. These sustained improvements might represent adoption of new standards of care that were driven by PBF and, once adopted, the incentive was no longer necessary to maintain a high level of quality.

Another study investigated the discontinuation of incentives within the VA. Hysong *et al.* (2011) used outpatient clinical performance measure data from VA's External Peer Review Program in the United States to investigate the mean time to stability of performance after changing status from being actively monitored (i.e., incentivized) to being passively monitored (i.e., no incentive) and vice versa. The study found that regardless of whether a measure was incentivized, all measures remained stable or improved over time. Quality did not deteriorate for any of the measures in which incentives were removed.

Similarly, Boland *et al.* (2010) found that a radiologist PBF program in the United States had a significant positive effect on the quality of care, measured by expediting final report turnaround times, which continued after its discontinuation.

In the United Kingdom, Kontopantelis *et al.* (2014) investigated the effect of discontinuing financial incentives for aspects of care for patients with asthma, coronary heart disease, diabetes, stroke and psychosis on eight clinical quality indicators withdrawn from a national incentive scheme. The study found that the level of performance achieved prior to incentive discontinuation was generally maintained, with some difference by indicator and disease condition.

However, Lester *et al.* (2010) found that discontinuing incentives was associated with a decrease in performance of screening for diabetic retinopathy and screening for cervical cancer in the United States.

Only one paper investigated the effect on health expenditure. Fiorentini *et al.* (2013) estimated the effect of discontinuing financial incentives offered to PHC providers in exchange for containing hospital expenditure in an Italian region. The study estimated the effect on hospital expenditure and found no significant effect on avoidable hospital

expenditure and total hospital expenditure. The evidence provided indicated that the discontinued incentives did not initially affect physicians' behavior.

To conclude, most the reviewed studies provided evidence that quality did not deteriorate for any of the measures in which incentives were removed. No significant effect was also observed on healthcare expenditure.

6.2.3.2 Introducing Provider Incentives

Evidence is lacking on the effects of discontinuing provider incentives in general and on the effects in low- and middle-income countries in particular. The available studies on provider incentives focused on the effects of introducing rather than discontinuing incentives to healthcare providers. Moreover, all studies retrieved were conducted in high-income settings. We did not find any study investigating the effect of discontinuing provider incentives in a low- or middle-income setting. Therefore, we complement our review in the previous subsection by reviewing the evidence on the effects of introducing provider incentives in an attempt to anticipate the likely effects of discontinuing these incentives in low- and middle-income countries.

Three systematic reviews of the literature are particularly relevant for the scope of this study: Das *et al.* (2016), Witter *et al.* (2012) and Eldridge & Palmer (2009). Das *et al.* (2016) assessed the existing evidence on the effects of PBF on the quality of maternal and child health care in low- and middle-income countries. The review found some evidence that PBF was associated with positive effects but only on the process quality of maternal and child health. This included adherence to standard protocols and guidelines for management of health conditions. The effects of PBF on delivery, emergency obstetric and neonatal care, postnatal care and under-five child care were not investigated in the studies included in this review. More importantly, Das *et al.* (2016) found weak evidence that PBF was associated with positive effects on maternal and neonatal health outcomes and out-of-pocket expenses. PBF was also found to have a few negative effects on structural quality.

Witter *et al.* (2012) assessed the existing evidence on the effects of PBF on the provision of health care and health outcomes in low- and middle-income countries. The review concluded that the current evidence base is too weak to draw general conclusions. The evidence provided suggests that the effects of PBF depend on the interaction of several variables. A key variable is the design of PBF (e.g., who receives

payments, the magnitude of the incentives, the targets and how they are measured). Other key variables include but are not limited to the amount of funding PBF receives, the strength of technical support PBF gets as well as contextual factors, such as the organizational context in which PBF is implemented.

Eldridge & Palmer (2009) assessed the existing evidence on PBF in low-income countries only. The review found significant weaknesses in the existing evidence base on the effectiveness of PBF initiatives. Eldridge & Palmer (2009) concluded that the lack of evidence on the effects of any type of PBF in any low-income country is mostly due to the absence of control groups.

Besides the systematic reviews, a quick review of studies on the effects of introducing provider incentives in low- and middle-income settings also suggests that the available evidence is mixed and of low quality. The evidence is particularly limited when we investigate the effects on patient health outcomes rather than process indicators. The health outcomes we focused on were those of family planning, and maternal and child health. A summary of these studies is provided in Table 6.2.b.

Effect on Family Planning. The effects of PBF on family planning outcomes are limited (see Table 6.2.b). Two of the three studies reviewed reported no effects (Soeters *et al.*, 2011; Falisse *et al.*, 2015). As per Soeters *et al.* (2011), PBF did not have significant effects on family planning output and patient knowledge indicators in Congo, specifically having heard about family planning and currently using a modern contraceptive method. Even when accompanied by removing user fees, Falisse *et al.* (2015) reported no effect of PBF in Burundi on fittings of intra-uterine devices (IUD) as a family planning method. However, mixed interventions in Egypt, including PBF, were associated with an increase in the current use of modern contraceptive methods (Grun & Ayala, 2006). Besides PBF, interventions in Egypt included quality improvement through facility accreditation and user fee introduction.

Effect on Maternal Health. A total of eight studies reported on the effects of PBF on maternal health. As per the evidence, PBF had variable effects on maternal health (see Table 6.2.b).

Table 6.2.b: Evidence on the effectiveness of introducing provider incentives

Study	Intervention	Outcome measure	Reported effect on outcome
Family planning			
Soeters *et al.* (2011)	-PBF	-Heard about family planning	None
		-Current use of modern contraception	None
Falisse *et al.* (2015)	-Demand: Cost of care removal -Supply: PBF	-Fitting of IUD as a family planning method	None
Grun & Ayala (2006)	-Demand: User fees -Supply: PBF, other	-Current use of modern contraception	Positive
Maternal health			
Basinga *et al.* (2011)	PBF	-Any ANC	None
		-At least 4 ANC visits	None
		-Institutional delivery	Positive
		-Anti-tetanus vaccination during ANC visit	Positive
Bonfrer *et al.* (2014)	PBF	-More than 1 ANC visit	None
		-First trimester ANC	None
		-Blood pressure measured during pregnancy	Positive
		-At least 1 anti-tetanus vaccination	Positive
		-Institutional delivery	None
Soeters *et al.* (2011)	PBF	-Institutional delivery	Positive
		-Episodes using modern health facility or pharmacy	Positive
		-ANC composite score 100% for 4 indicators	None
Van de Poel *et al.* (2016)	PBF	-At least 2 ANC visits	None
		-Institutional delivery	Positive
Falisse *et al.* (2015)	-Demand: Cost of care removal -Supply: PBF	-ANC visits	Positive
		-Anti-tetanus vaccination of pregnant women	Positive
		-Institutional delivery	None
		-Postnatal visits	None
		-Malaria visits	None

Study	Intervention	Outcome measure	Reported effect on outcome
Gertler *et al.* (2014)	-Demand: Health insurance -Supply: PBF	-Number of ANC visits - Anti-tetanus vaccination -C-section	Positive Positive Positive
Grun & Ayala (2006)	-Demand: User fees -Supply: PBF, other	-At least 4 ANC visits -Skilled-assisted delivery	None None
Nguyen *et al.* (2012)	-Demand: Vouchers, cost of care removal -Supply: PBF	-Skilled-assisted delivery -Institutional delivery -C-section -Any ANC check-up -At least 3 ANC check-ups -At least 1 ANC check-up with a qualified provider -At least 1 postnatal check-up with a qualified provider	Positive Positive None Positive Positive Positive Positive
Child health			
Basinga *et al.* (2011)	-PBF	-Younger than 23 months preventive visit, previous 4 weeks -24–59 months preventive visit, previous 4 weeks -12–23 months fully vaccinated	Positive Positive None
Bonfrer *et al.* (2014)	-PBF	-Child fully vaccinated at 1 year	Positive
Peabody *et al.* (2014)	-PBF	-Acute ongoing infection measured by C-reactive protein -Anemia measured by blood hemoglobin levels -Wasted children (underweight for height) -Parental assessment of improvement in children's health using a general self-reported health measure	None None Positive Positive
Skiles *et al.* (2015)	-PBF	-Reported illness (diarrhea, fever and/or ARI) -Facility care-seeking (diarrhea, fever and/or ARI) -Reported illness (diarrhea and/or fever) -Facility care-seeking (diarrhea and/or fever) -Treatment received diarrhea and/or fever	None None None None Positive
Soeters *et al.* (2011)	-PBF	-Vaccination composite score (for children under 1 year) 100% for 4 indicators	None

Study	Intervention	Outcome measure	Reported effect on outcome
Van de Poel *et al.* (2016)	-PBF	-12–24 months fully vaccinated	None
		-Neonatal mortality	None
Falisse *et al.* (2015)	-Demand: Cost of care removal	-Perinatal deaths/deliveries	None
	-Supply: PBF	-Anti-polio vaccination	None
		-Anti-BCG vaccination	None
		-Anti-MMR vaccination	None
		-Anti-DPT vaccination	None
Gertler *et al.*, 2014	Demand: Health insurance	-Birth weight (in grams)	Positive
	Supply: PBF	-Low birth weight (<2,500 grams)	Positive
		-Neonatal mortality per 1,000 live births	Positive
Grun & Ayala (2006)	-Supply: Local personnel capacities' reinforcement	-Diarrhea mortality rate	None
	-Supply: Health infrastructure enhancement	-Diarrhea mortality rate	None
	-Demand: User fees	-Diarrhea mortality rate	None
	-Mixed (demand: user fees supply: PBF, other)	-Vaccination coverage (measles)	Positive
		-Use of medical treatment for child fever/cough cases	Positive
		-Share of diarrhea cases in children receiving medical treatment	None

IUD: Intra-uterine device. C-section: Cesarean section. ARI: Acute respiratory infection.

Out of the eight studies, seven reported on ANC measured by the number of visits. Three out of the seven studies reported positive effects of PBF on the number of ANC visits in Burundi, Argentina and Bangladesh (Falisse *et al.*, 2015; Gertler *et al.*, 2014; Nguyen *et al.*, 2012). However, none of the three studies investigated the effect of PBF as a stand-alone policy. In all three studies, PBF was introduced together with other demand-side interventions that involved removing the cost of care. The remainder of the studies reporting on ANC found no effect on the number of visits in Rwanda, Burundi, Cambodia and Egypt (Basinga *et al.*, 2011; Bonfrer *et al.*, 2014; Van de Poel *et al.*, 2016; Grun & Ayala, 2006).

Despite the limited effect of PBF on the number of ANC visits, evidence on the positive effect of PBF on the quality of ANC is more consistent. All four studies reporting on the quality of ANC found a positive effect of PBF on the quality of ANC. In Rwanda, PBF significantly improved the quality of care measured by anti-tetanus vaccination during ANC and standardized total quality scores. This improvement was not, however, associated with increase in the use of care measured by the probability that women receive any ANC or that they have at least four ANC visits (Basinga *et al.*, 2011). Similar effects were captured by Bonfrer *et al.* (2014) in Burundi, where the quality of care provided during ANC visits improved significantly as a result of PBF despite the fact that the number and timeliness of visits did not change. The quality of ANC was captured by both blood pressure (BP) measurement and anti-tetanus vaccination. Falisse *et al.* (2015) and Gertler *et al.* (2014) also reported a significant positive effect of PBF in Burundi and Argentina, respectively, on anti-tetanus vaccination

With respect to delivery care, four out of the six studies reporting on institutional delivery found significant positive effects. Basinga *et al.* (2011) captured a significant increase in the probability of institutional delivery associated with PBF in Rwanda. The PBF scheme in Congo was also associated with a significant increase in institutional delivery (Soeters *et al.*, 2011). Similarly, Van de Poel *et al.* (2016) captured a significant increase in Cambodia in the proportion of births occurring in incentivized public health facilities. However, the effects on delivery in public facilities were much greater if PBF was accompanied by maternity vouchers that cover user fees. No significant effects are observed among the poorest women though. Nguyen *et al.* (2012) provided evidence that combining cash incentives for individuals

(i.e., vouchers) with cash incentives for healthcare providers (i.e., PBF) could significantly increase institutional delivery.

Effect on Child Health. A total of nine studies reported on the effects of PBF on child health. Table 6.2.b provides evidence that PBF had limited effects, in general, on child health outcomes.

Only two out of the seven studies that report on the effects of vaccination captures significant positive effects of PBF. As per Bonfrer *et al.* (2014), PBF significantly increased the probability of a child being fully vaccinated in Burundi, with effects being more pronounced among the poor. In Egypt, PBF accompanied by other interventions significantly increased the measles vaccination rate (Grun & Ayala, 2006). However, Basinga *et al.* (2011), Soeters *et al.* (2011), Van de Poel *et al.* (2016), Falisse *et al.* (2015) and Gertler *et al.* (2014) reported no effects of PBF on vaccination rates in Rwanda, Congo, Cambodia, Burundi and Argentina, respectively.

Evidence on child mortality is limited. Only three studies reported on the effects of PBF on child mortality, one out of which captured significant positive effects in Argentina (Gertler *et al.*, 2014). The study provided evidence that combining health insurance with PBF had significant positive effects on child health, where beneficiaries had lower chance of in-hospital neonatal mortality. Approximately half of this reduction resulted from preventing low birth weight and half from better postnatal care. In parallel, Van de Poel *et al.* (2016) and Falisse *et al.* (2015) did not find any effects of PBF on neonatal mortality and perinatal mortality, respectively, in Cambodia and Burundi.

With respect to medical treatment, Skiles *et al.* (2015) reported a significant positive effect of PBF on treatment received by children with diarrhea or fever at health facilities in Rwanda. Similarly, mixed interventions in Egypt, including PBF, significantly improved child access to medical treatment (fever/cough). However, we did not observe a significant effect on the share of diarrhea cases in children receiving medical treatment (Grun & Ayala, 2006).

To conclude, we find that the evidence on the effectiveness of introducing PBF in improving family planning, maternal health and child health outcomes is mixed. This is to some extent because the PBF schemes introduced in low- and middle-income countries are mixed as well. The most consistent evidence found on the effectiveness

of PBF is with respect to the quality of ANC and medical treatment among children. These are the areas that we expect to be affected the most if provider incentives are discontinued.

6.3 ECONOMETRIC STRATEGY

Since the publication of the seminal article Rajan & Zingales (1998), the estimation of models with interaction effects has been applied widely (e.g., Castro *et al.*, 2004; Easterly *et al.*, 2004; Spilimbergo, 2009). To capture the effects of discontinuing provider incentives, we propose a simple method for modeling interactions between "being contracted" and a time dummy for year 2014 in which PBF is not in operation. The 2005, 2008 and 2014 waves of the Egypt Demographic and Health Survey (DHS) are combined in a panel. GLS RE and FE (within) regressions are then used to estimate the effects of being contracted in 2014 versus being contracted during the period 2005-2008, when PBF was still in operation.

Provider incentives were only offered in contracted facilities. Our hypothesis is that discontinuing PBF in 2008 affects the health outcomes of these facilities. We include two subgroups in the analysis: facilities that are non-contracted in 2005 and continue to be so in 2008 and 2014 (subgroup 1) and facilities that are contracted in 2005 and continue to be so in 2008 and 2014 (subgroup 2). We expect only the latter subgroup to be affected by the discontinuation of PBF by the end of 2008 as provider incentives were never offered in subgroup (1) of facilities.

To test our hypothesis, we would use either of the two main estimation methods of panel data models: the RE model and the FE model. The main distinction between both models is whether the unobserved facility effect has elements that are correlated with the explanatory variables in the model, not whether these effects are stochastic or not (Greene, 2008). When using RE, we assume that the variation across facilities is random and not correlated with the explanatory variables included in the model. If this is the case, time-invariant variables could play a role as explanatory variables. The RE model could be estimated using GLS as follows:

$$y_{it} = \beta\, cont_{it} + \gamma\, d_{2014} + \delta\, cont_{it} * d_{2014} + \zeta\, fac_i + \eta\, dist_i + \alpha + u_i + \varepsilon_{it} \quad (6.7)$$

The dependent variable y_{it} denotes a health outcome of interest y for facility i at time t. t = 0, 1 and 2 for years 2005, 2008 and 2014, respectively. The variable $cont_{it}$ = 1 if the facility is contracted in all three years and $cont_{it}$ = 0 if the facility is non-contracted in all three years. The coefficient β captures permanent differences between contracted and non-contracted facilities. The variable d_{2014} = 1 for year 2014 and d_{2014} = 0 for years 2005 or 2008. The interaction term $(cont_{it} * d_{2014})$ equals one for contracted facilities in 2014 and zero otherwise. In this context, the coefficient of the interaction term δ captures the effects of being contracted on a health outcome of interest in year 2014 compared to being contracted during the period 2005-2008. For outcomes, such as *mcp* and *ancprov*, a negative sign of the coefficient δ indicates that positive effects of being contracted weaken in 2014 versus 2005-2008. We could, therefore, interpret this sign as a negative effect of discontinuing PBF. For outcomes, such as *childmort*, a positive rather than a negative sign of δ is interpreted as a negative effect of discontinuing PBF. The terms fac_i and $dist_i$ represent facility-level controls and district-level controls, respectively. u_i is the between-facility error (i.e., a group-specific random element). ε_{it} is the within-facility error.

For each health outcome, we report the results from two specifications of equation (6.7). Our model's specification (1) includes facility-level controls only. Specification (2) includes both facility- and district-level controls. The standard errors in all regressions are robust.

For some health outcomes, however, the RE model is not considered appropriate. This is the case if facilities are likely to have some characteristics that make them more likely to participate in the treatment or to benefit or not benefit from it. In this context, the FE model is preferred. When using FE, we assume that some characteristics within the facility itself could affect or bias the explanatory or dependent variables. To control for this, FE models remove the effects of these time-invariant characteristics to enable us to capture the net effects of the explanatory variables on the dependent variable. When using FE, we also assume that those characteristics are unique to the facility and are not correlated with other facility characteristics.

Our FE model is formulated as follows:

$$y_{it} = \beta\, cont_{it} + \gamma\, d_{2014} + \delta\, cont_{it} * d_{2014} + \zeta\, fac_i + \eta\, dist_i + \alpha_i + \varepsilon_{it} \quad (6.8)$$

In equation (6.8), α_i is the unknown intercept for each facility *i*. The FE approach regards α_i as a group-specific constant term in the regression model. ε_{it} is the error term. The time-invariant characteristics are absorbed by the intercept. The standard errors in all regressions are robust.

Considerations for Model Choice. Before any discussion of the estimated effects, a decision should be made on which of the two models provides efficient and consistent estimates of parameters. To decide on this, we employ two approaches to test for RE versus FE. The first approach is based on the Hausman test (Hausman, 1978). The second approach is based on a test of over-identifying restrictions that uses an artificial regression approach (Arellano, 1993; Wooldridge, 2002). Both test statistics are compared against a chi-squared distribution.

The Hausman test checks whether the error term is correlated with the explanatory variables in our models. The null hypothesis of the test is that the error is not. On one hand, it is reasonable to run the FE model given our panel data structure as FE yields more consistent estimates. On the other hand, whenever it is statistically plausible, the RE model is preferred to the FE model as the former yields more efficient estimates. Although we cannot run the Hausman test when robust standard errors are specified, we could still run the generalized test by running the models without specifying robust errors.

The second approach we employ is based on a test of over-identifying restrictions. Both FE and RE models impose additional orthogonality conditions that are basically over-identifying restrictions. RE models imply additional orthogonality conditions that the explanatory variables are not correlated with the facility-specific error (RE). FE models imply additional orthogonality conditions that the explanatory variables are not correlated with the idiosyncratic error. The extra orthogonality conditions that cause the RE estimates to be more efficient than the FE estimates could be considered as over-identifying restrictions that could be tested.

To run the test, we follow the artificial regression approach explained by Arellano (1993) and Wooldridge (2002). We re-estimate our RE models, but this time, in addition to our original explanatory variables, we include transformations of these variables into deviation-from-mean form. We subsequently test for the significance of these additional variables. Rejection of the null hypothesis implies that the FE model is preferred to the RE model.

In this context, it is important to note that the test of over-identifying restrictions is better than the Hausman test as the former enables us to test RE against FE when the errors are heteroskedastic or have intragroup correlation. Tests of over-identifying restrictions extend directly to heteroskedastic- and cluster-robust versions.

6.4 DATA

We use three waves of the Egypt DHS (2005, 2008 and 2014) to calculate our health outcomes of interest at the facility level. Information obtained from Egypt's MOH is used to capture interventions at the facility level and calculate facility-level controls. We also use data from Egypt's Central Agency for Public Mobilization and Statistics (CAPMAS) to calculate a set of district-level social and economic controls. A number of regional dummies is included as well.

Dependent Variables. The dependent variables included in this chapter are similar to some of the family planning, maternal health and child health outcomes included in the analyses of chapters 3, 4 and 5. However, in this chapter, we focus on the health outcomes that reflect the health services targeted by the PBF scheme in Egypt as well as the quality of these services. A list of these outcomes together with a description of how they are linked to the PBF scheme is provided in Table 6.3.

To construct these outcomes, we follow the same steps discussed in details in chapter 3. First, we use Quantum GIS 2.8.2 to spatially link women interviewed in each of the Egypt DHS waves to their nearest mapped facilities. Second, we recode and compute health outcomes at the facility-level using Stata 12.0. We, finally, combine the 2005, 2008 and 2014 facility-level outcomes in a panel.

Table 6.3: Description of dependent variables and their link to PBF scheme

	Outcome	Description*	Relevant PBF scheme indicator	Link to scheme
Family planning	*mcp*	Proportion of women currently using any modern contraceptive method	Number of new users of all types of modern contraceptive methods among married women of reproductive age in the catchment area	Direct
	contsid	Proportion of current users of selected contraceptive methods informed of side effect of the method used	Number of new users of all types of modern contraceptive methods among married women of reproductive age in the catchment area	Indirect (quality of service)
	contoth	Proportion of current users of selected contraceptive methods informed of other methods of contraception that could be used	Number of new users of all types of modern contraceptive methods among married women of reproductive age in the catchment area	Indirect (quality of service)
ANC	*ancprov*	Proportion of women attended for ANC by skilled health personnel	Number of pregnant women receiving regular ANC visits compared to the total number of pregnant women in the catchment area	Direct
	anc4	Proportion of women who received four or more ANC visits	Number of pregnant women receiving regular ANC visits compared to the total number of pregnant women in the catchment area	Direct
	anciron	Proportion of women who received iron supplements as an ANC component	Number of pregnant women receiving regular ANC visits compared to the total number of pregnant women in the catchment area	Indirect (quality of service)
Child health status	*childmort*	Proportion of deaths at age 0–5 years to live-born children	Number of children fully vaccinated in the catchment area	Direct (proxy measure)

*Definitions are obtained from the World Health Organization (WHO).

Explanatory Variables. Our main explanatory variable is $cont_{it}$, which is a dummy variable that switches on only when a health facility is contracted in all three years 2005, 2008 and 2014. $cont_{it}$ draws on information from Egypt's MOH on whether a facility is contracted. Our dataset includes facilities that are contracted in all three years 2005, 2008 and 2014 in addition to facilities that are non-contracted in 2005 and continue to be so in both years 2008 and 2014. Accordingly, any changes in the outcomes of non-contracted facilities cannot be attributed to discontinuing provider incentives. We remove facilities that witness any alteration in their contracting status between 2005 and 2014 from the dataset.

In all regression models, we include facility-level characteristics, district-level social and economic characteristics as well as regional dummies to control for any potential discrepancies at facility, district and regional levels, respectively. A detailed list of all the control variables included in our analyses is provided in chapter 3 (see Table 3.4).

6.5 RESULTS

In this section, we present the descriptive statistics of the subsample used and discuss the estimated effects of discontinuing provider incentives on family planning and maternal and child health.

6.5.1 Descriptive Statistics

Table 6.4 summarizes the descriptive statistics of the facility-level characteristics, the district-level characteristics as well as the health outcomes of our subsample. We observe moderate variation in labor force, infrastructure (*inf*) and coverage (*cov*) across Egypt's health facilities. Moreover, we observe moderate variation in the socio-economic profiles across districts in Egypt. However, relatively large standard deviations are reported with respect to the health outcomes of facilities included in our analyses.

Tables 6.5 and 6.6 compare facility and district characteristics, respectively, by group of facilities. On average, the characteristics of contracted facilities are more favorable at both facility- and district- levels compared to other facilities. The most significant differences observed at the facility level are with respect to the number of practitioners (*pract*), specialists (*spec*) and pharmacists (*pharm*) who work in facilities (Table 6.5).

Table 6.4: Descriptive statistics

Variable	N	Mean	Standard deviation	Minimum	Maximum
Facility characteristics	2,808	3.533	4.681	0.000	37.000
Practitioners	2,307	0.787	2.070	0.000	22.000
Specialists	2,739	4.090	5.628	0.000	40.000
Pharmacists	2,856	10.875	10.869	0.000	99.000
Nurses	2,655	1.534	1.862	0.000	16.000
Lab technicians	2,286	0.252	0.952	0.000	14.000
X-ray technicians	2,730	1.481	1.797	0.000	17.000
Health observers	2,274	0.420	0.925	0.000	8.000
Social workers	2,499	1.540	0.699	0.000	2.000
Building condition	2,610	31.092	57.262	0.290	460.035
Population coverage					
District characteristics					
Illiteracy	3,336	32.066	10.748	0.000	56.390
Unemployment	3,336	9.714	4.813	0.000	23.550
Income dependency	3,336	6.337	17.401	0.000	87.402
Inaccessibility to electricity	3,336	1.848	5.728	0.000	69.092
Inaccessibility to potable water	3,336	5.338	9.232	0.000	90.204
Family size	3,336	4.324	0.391	2.580	6.170
HH overcrowding	3,336	1.141	0.109	0.840	1.930
Population size	3,336	31.185	17.152	0.005	117.380
Health outcomes					
Family planning					
Modern contraceptive prevalence	1,508	50.396	17.373	0.000	100.000
Knowledge of side effects	1,253	46.990	23.484	0.000	100.000
Knowledge of contraceptives	1,254	55.276	22.817	0.000	100.000
ANC					
ANC by skilled health personnel	1,267	77.741	21.881	0.000	100.000
4+ visits	1,267	68.352	24.473	0.000	100.000
Iron supplementation	1,267	51.488	25.502	0.000	100.000
Child health status					
Under-5 mortality	1,367	2.683	5.042	0.000	31.250

N denotes the number of observations.

Table 6.5: Two-sample t-test of facility characteristics of contracted and non-contracted facilities

	Non-contracted	Contracted	Difference
Practitioners	3.337	5.750	-2.413***
			(0.555)
Specialists	0.714	1.650	-0.936***
			(0.277)
Pharmacists	3.818	7.039	-3.221***
			(0.662)
Nurses	10.987	9.597	1.390
			(1.292)
Lab technicians	1.481	2.139	-0.658***
			(0.228)
X-ray technicians	0.248	0.304	-0.056
			(0.132)
Health observers	1.421	2.130	-0.709***
			(0.213)
Social workers	0.420	0.410	0.011
			(0.124)
Building condition	1.509	1.875	-0.366***
			(0.085)
Population coverage	31.845	23.113	8.732
			(6.917)

Standard errors are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

Table 6.6: Two-sample t-test of district characteristics of contracted and non-contracted facilities

	Non-contracted	Contracted	Difference
Illiteracy	32.186	30.476	1.710
			(1.262)
Unemployment	9.788	8.731	1.057*
			(0.565)
Income dependency	6.524	3.849	2.675
			(2.043)
Inaccessibility to electricity	1.928	0.786	1.142*
			(0.672)
Inaccessibility to potable water	5.462	3.701	1.761
			(1.084)
Family size	4.328	4.263	0.065
			(0.046)
HH overcrowding	1.139	1.170	-0.031**
			(0.013)
Population size	30.806	36.200	-5.394***
			(2.009)

Standard errors are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

Table 6.5 also shows that the building condition of contracted facilities is significantly better than that of non-contracted facilities. At the district level, significant differences are observed with respect to four out of eight socio-economic indicators. These four indicators are unemployment (*unemp*), inaccessibility to electricity (*elect*), HH overcrowding (*crowd*) and the population size (*pop*) (Table 6.6).

Table 6.7 presents the means of health outcomes of groups of health facilities by year. We find that six out of seven outcomes improved between 2005 and 2008 for non-contracted facilities. However, three out of seven outcomes improved for contracted facilities during the same period. Between years 2008 and 2014, we observe that six out of seven outcomes improved for all types of facilities. We also observe that some outcomes, such as *mcp* and *contoth*, are steadily improving over time for all types of facilities.

As previously noted, financial incentives to contracted healthcare providers were discontinued only by the end of 2008. Hence, we expect the change in outcomes of contracted facilities between 2008 and 2014 to reflect this intervention. The last two columns of Table 6.7 show that only one out of seven outcomes deteriorated during the period 2008-2014. Although most the outcomes of contracted facilities improved after discontinuing PBF in 2008, it is important to note that the magnitude of this improvement is less than that achieved by non-contracted facilities for some outcomes. For example, while knowledge of other contraceptive methods (*contoth*) increased by one percentage points (ppts) between 2008 and 2014 among women with access to contracted facilities, *contoth* increased by more than five ppts among women with access to non-contracted facilities (Table 6.7).

6.5.2 Estimated Effects of Discontinuing Provider Incentives

A decision should be made on which of the RE/FE models provides efficient and consistent estimates of parameters before any discussion of the estimated effects. To decide on this, two approaches are used as discussed in section 6.3: the Hausman test and the test of over-identifying restrictions. We use our preferred model's specification that includes all controls. Our subsample includes health facilities that are contracted in all three years 2005, 2008 and 2014 in addition to facilities that are non-contracted in 2005 and continue to be so in both 2008 and 2014.

Table 6.7: Health outcomes by year

	Non-contracted			Contracted		
	2005	2008	2014	2005	2008	2014
Modern contraceptive prevalence	52.096 (0.778)	52.286 (0.825)	54.720 (0.721)	54.655 (2.198)	58.711 (3.443)	61.250 (2.376)
Knowledge of side effects	46.367 (1.281)	47.921 (1.292)	46.514 (1.231)	44.979 (5.351)	40.429 (4.037)	52.458 (3.697)
Knowledge of contraceptives	53.435 (1.158)	54.234 (1.390)	57.140 (1.216)	53.207 (3.263)	58.345 (4.102)	59.018 (3.266)
ANC by skilled health personnel	71.106 (1.186)	72.680 (1.209)	91.962 (0.561)	78.830 (4.140)	63.180 (6.003)	94.725 (1.981)
4+ visits	59.803 (1.269)	63.658 (1.379)	84.747 (0.765)	63.839 (4.182)	55.952 (5.643)	87.593 (2.346)
Iron supplementation	48.357 (1.202)	36.857 (1.354)	67.116 (1.070)	60.470 (4.128)	43.608 (4.937)	81.544 (3.107)
Under-5 mortality	3.628 (0.287)	2.380 (0.280)	2.248 (0.224)	3.102 (1.265)	1.484 (0.615)	1.763 (0.810)

Standard errors are reported in parentheses.

For equations (6.7) and (6.8), Table 6.8 reports the FE (within) and the RE GLS estimates together with the results of Hausman's and Sargan-Hansen's tests.[1] As per the results of the Hausman test, we reject the null hypothesis of no endogeneity for three out of seven health outcomes. The FE model should be used to estimate *ancprov, anc4* and *anciron*. The Sargan-Hansen Chi-square test statistics indicate that the FE model is preferred to the RE model for four of our outcomes: *ancprov*, *anc4*, *anciron* and *childmort*. This conclusion coincides with the findings of the Hausman test for all outcomes except *childmort*. Only for *contoth*, the results of our tests of over-identifying restrictions indicate that the RE estimates are equivalent to the pooled ordinary least-squares (POLS) estimates.

Table 6.9 presents the results of the appropriate model specification for each health outcome. We report the results of two specifications. The model's specification (1) includes facility-level controls only; specification (2) includes both facility- and district-level controls and is our preferred. Our estimates capture the effects of being contracted in year 2014, after PBF was discontinued, versus being contracted during the period 2005-2008, when PBF was still in operation.

[1] We perform the Sargan-Hansen test using the Stata command *xtoverid* developed by Schaffer and Stillman (2010).

Table 6.8: Estimated differential effects of discontinuing provider incentives

	Outcome	D_2014=1*D_CONT=1		Hausman Chi-sq	Sargan-Hansen Chi-sq
		FE	RE		
Family planning	Modern contraceptive prevalence	3.129 (4.613)	0.027 (4.123)	2.39	2.455
	Knowledge of side effects	-3.271 (7.976)	0.341 (6.716)	0.31	0.463
	Knowledge of contraceptives	-13.802** (5.752)	-15.145*** (4.677)	0.04	-
ANC	ANC by skilled health personnel	-12.043*** (4.406)	-6.743* (3.672)	7.62**	10.091***
	4+ visits	-4.800 (6.245)	-4.134 (5.057)	6.57**	8.018**
	Iron supplementation	-19.295* (10.976)	-6.062 (7.261)	8.36**	9.508***
Child health status	Under-5 mortality	1.719*** (0.654)	-0.204 (0.806)	3.57	6.369**

Each row represents a separate regression. Robust standard errors across clusters are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

Table 6.9: Estimates of the most appropriate model specification for discontinuing provider incentives

	Outcome	D_2014=1*D_CONT=1	
		(1)	(2)
Family planning	Modern contraceptive prevalence	-0.567 4.455	0.027 (4.123)
	Knowledge of side effects	-0.342 6.781	0.341 (6.716)
	Knowledge of contraceptives	-14.260*** 4.834	-15.145*** (4.677)
ANC	ANC by skilled health personnel	-12.043*** (4.406)	-12.043*** (4.406)
	4+ visits	-4.800 (6.245)	-4.800 (6.245)
	Iron supplementation	-19.295* (10.976)	-19.295* (10.976)
Child health status	Under-5 mortality	1.719*** (0.654)	1.719*** (0.654)

Each row represents a separate regression. Robust standard errors across clusters are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

We find that discontinuing provider incentives is associated with multiple negative effects on the outcomes of contracted facilities (Table 6.9). With respect to family planning, a significant negative effect is captured with respect to knowledge of other contraceptive methods that could be used (*contoth*). Table 6.9 indicates that the proportion of women with access to contracted facilities, who are informed of other contraceptives *(contoth)*, decreased in 2014 by more than 15 ppts compared to the period 2005-2008.

In parallel, we find that discontinuing provider incentives is associated with a significant negative effect on the likelihood of receiving ANC by skilled health personnel (*ancprov*) (Table 6.9). We realize that *ancprov* is an outcome that is directly and strongly related to provider incentives. The choice of a woman to use the skilled care of a healthcare provider is expected to be influenced by the quality of care that the provider offers. In this sense, if discontinuing financial incentives demotivates providers from offering the same quality of care, women could become less encouraged to be attended for ANC by skilled health personnel.

We also find that discontinuing provider incentives is associated with a significant negative effect on the likelihood of receiving iron supplements during pregnancy *(anciron)*. Table 6.9 shows that the proportion of women with access to contracted facilities, who receive iron supplements *(anciron)*, decreased by about one fifth in 2014 compared to the period 2005-2008. This finding suggests that discontinuing provider incentives worsens the quality of ANC provided in contracted facilities.

More importantly, Table 6.9 provides evidence that discontinuing provider incentives has a significant negative effect on child mortality. We observe that discontinuing provider incentives is associated with a significant increase by 2 ppts in under-five mortality (*childmort*) among children with access to contracted facilities. This negative effect suggests that discontinuing provider incentives is associated with lower quality of child health services provided, which deter women from its use.

As previously noted, one of the indicators based upon which a contracted facility qualifies for financial incentives is the protection provided by family planning services denoted by the years of protection. We observe that discontinuing provider incentives has an insignificant effect on modern contraceptive prevalence (*mcp*) among women with access to contracted facilities in 2014 (Table 6.9). Another

indicator is ANC visits per pregnant woman. We observe that discontinuing the incentives has a negative effect on ANC coverage (at least four visits) *(anc4)*. However, the effect observed is not significant (Table 6.9).

6.6 ROBUSTNESS CHECKS

As noted earlier, the RE and FE models are used to estimate the effects of discontinuing provider incentives. We perform a number of checks to test for facility FE and RE. These models are tested against the POLS model.

Facility-FE Model. First, we test the FE model against the POLS model. To do so, we use the F-test that jointly tests that all facility effects are zero for each of the health outcomes. We do not report correct cluster-robust Huber/White standard errors. The second column of Table 6.10 reports the F-test statistics for FE. We reject the null hypothesis that facility effects equal zero for most of our outcomes. Thus, the FE model is preferred to the POLS model as significant facility FE are observed. If we choose to use the POLS model, unobserved facility heterogeneity will cause our estimates to be biased and inconsistent.

Table 6.10: Summary of results of diagnostic tests

	Facility FE: F statistic	RE: Chi-sq statistic	RE vs. FE: Hausman's Chi-sq statistic	RE vs. FE: Sargan-Hansen Chi-sq statistic	Preferred model
Modern contraceptive prevalence	1.98***	7.15***	2.39	2.455	RE
Knowledge of side effects	1.10	0.38	0.31	0.463	RE
Knowledge of contraceptives	1.02	0.00	0.04	-	POLS/RE
ANC by skilled health personnel	2.29***	1.22	7.62**	10.091***	FE
4+ visits	2.18***	1.00	6.57**	8.018**	FE
Iron supplementation	1.22*	5.84***	8.36**	9.508***	FE
Under-5 mortality	1.27**	1.51*	3.57	6.369**	FE

RE Model. Second, we test the RE model against the POLS model. To do so, we use the Breusch and Pagan Lagrangian multiplier (LM) test. The LM test statistic is compared against a chi-squared distribution. Correct cluster-robust Huber/White

standard errors are reported. The third column of Table 6.10 reports the chi-sq test statistics for RE. We reject the null hypothesis that the variance of the unobserved FE is zero for *mcp*, *anciron* and *childmort*. Significant RE are observed in the panel data used to estimate some outcomes. For these outcomes, the RE model is preferred to the POLS model.

RE versus FE Model. As previously discussed, the Hausman test and a test of over-identifying restrictions are used to test for RE versus FE. For *contoth*, the POLS model is preferred to the RE or the FE model. However, we prefer to use a panel data model to exploit the panel structure of our data. So, we would rather use the RE or the FE model. We could still report the results that the POLS model yields, whenever it is preferred, to support the robustness of our results yielded by the panel data model we choose (RE or FE).

The Hausman and the Sargan-Hansen Chi-square tests' statistics are reported in the fourth and the fifth columns, respectively, of Table 6.10. These results support the findings of our tests for facility-FE and RE reported in the second and third columns, respectively, of Table 6.10. The preferred model for each of our health outcomes of interest is indicated in the last column of Table 6.10.

6.7 CONCLUSION

This chapter introduces evidence for the first time on the effect of discontinuing provider incentives in low- and middle-income countries, and highlights the effects of such discontinuation on the health services initially targeted by the PBF scheme. Data from three waves of the Egypt DHS (2005, 2008 and 2014) are collapsed at the facility level and RE/FE models are used in this context.

With respect to family planning, we find that discontinuing provider incentives had a significant negative effect on knowledge of other contraceptive methods that could be used. As for ANC, discontinuing the incentives had a significant negative effect on the likelihood of receiving ANC by skilled health personnel and receiving iron supplements during pregnancy. More importantly, we find that discontinuing the incentives was associated with higher under-five child mortality.

The findings of this chapter suggest that PBF schemes need to be applied carefully in low- and middle-income countries as negative effects are observed when provider

incentives are discontinued. The sustainability of funding remains a key challenge for continuing the incentives in most of these countries. Thus, adequate revenue streams should be generated to finance PBF schemes in order to avoid any potential negative effects of discontinuation. In addition, consideration of alternative health sector interventions is initially required before introducing any incentives in low-resource settings, especially in light of the weak evidence on the maintenance of performance levels after incentive discontinuation.

7. CONCLUSION

7.1 INTRODUCTION

This thesis investigates the effects of different interventions under Egypt's Health Sector Reform Program (HSRP) on a complete set of family planning, maternal health and child health outcomes during the period 2000-2014. To provide reliable evidence, a number of methods are used complementarily: difference-in-differences (DD), DD propensity score matching (PSM), fixed effects (FE), random effects (RE) and pooled ordinary least-squares (POLS). Four waves of the Egypt's Demographic and Health Survey (DHS) are collapsed at the facility level to calculate our health outcomes of interest. These outcomes are combined with information on different interventions at the facility level, facility-level characteristics, and district-level social and economic characteristics.

This chapter concludes the thesis. In section 7.2, we summarize the results obtained from the preceding empirical chapters. We derive some policy implications based on our empirical results in section 7.3. Section 7.4 outlines the limitations of the thesis. Finally, we propose some directions for future research in section 7.5.

7.2 SUMMARY OF FINDINGS

In this section, we summarize the results on the effects on family planning and maternal and child health obtained from the empirical chapters 3 to 6. The main contributions of the thesis are highlighted in accordance.

Chapter 3 that estimates the effect of improving the quality of health care through a facility accreditation program provides evidence that quality improvements can have multiple significant positive effects, especially on delivery care and child morbidity prevalence. However, for these effects to be sustained, a high level of commitment from the central government is indispensable. The chapter also provides evidence that facility accreditation as a stand-alone policy is ineffective in improving the utilization as well as the quality of ANC services.

In chapter 4, we found that introducing user fees drove demand for health services in two opposite directions. With respect to ANC, the positive effect of increased willingness to pay for an improved quality of service outweighed the negative effect of the price elasticity of demand. Introducing user fees was associated with a higher likelihood of receiving ANC by skilled health personnel, a higher likelihood of receiving at least four ANC visits and a higher likelihood of receiving iron supplements during pregnancy. However, the two effects offset each other with respect to the outcomes that reflect the utilization of family planning and delivery care services, women's access to health care, and child health status. No net effect at all was observed on these outcomes.

Again, we observe positive effects on both the utilization and the quality of ANC services when user fees are accompanied by two quality improvement interventions (facility accreditation and PBF of providers) (chapter 5). More notably, a positive effect on access to care was observed during our first study period (2000-2005). However, these effects were reversed during the second study period (2005-2008). In this context, it is important to note that user fees were introduced in 2004 and, therefore, the second study period is more likely to reflect the effects of user fees while the first period is more likely to reflect the effects of quality improvements. The positive effects of introducing user fees captured in chapter 4 could be partially driven by higher level of public awareness of the pro-poor exemption policy during the study period of this chapter (2008-2014) compared to the study period of chapter 5 (2005-2008).

It is also important to note that the positive effects reported in both chapters 4 and 5 were mainly with respect to ANC outcomes. No effects were reported with respect to the outcomes that reflect the utilization of family planning and delivery care services, and child health status. The findings of both chapters suggest that, even when accompanied by quality improvements, introducing user fees in low- and middle-income settings can have negative effects on access to and utilization of health care, especially in light of unpublicized exemptions and unsustainable quality improvements.

In chapter 6, we provide the first evidence on the effect of discontinuing provider incentives in low- and middle-income countries. We benefit from the fact that PBF

was replaced by “fee-for-service” in Egypt in 2008 to separately estimate the effect of incentive discontinuation on health outcomes that reflect the health services targeted by the PBF scheme as well as the quality of these services during the period 2008-2014. We found that discontinuing the incentives had a significant negative effect on four out of the seven considered health outcomes: knowledge of contraceptive methods, receiving ANC by skilled health personnel, receiving iron supplements during pregnancy and, more importantly, under-five child mortality. These multiple negative effects suggest that PBF schemes need to be applied carefully in low- and middle-income countries.

7.3 POLICY IMPLICATIONS

This thesis has important policy implications for improving family planning, maternal health and child health in low- and middle-income countries. First, improving the quality of care through facility accreditation can be particularly effective in improving delivery care and child health. If needed, accreditation can be accompanied by some interventions in order to meet equipment quality standards and strengthen staff’s competence in addressing maternal and child health needs. Second, the sustainability of the positive effects of quality improvements partially depends on the level of commitment from the central government.

Third, user fees are ineffective, in general, as a stand-alone policy. Introducing user fees should be part of a broader package of interventions that include addressing the quality of care in order to offset reduction in care utilization. Fourth, introducing user fees on the demand side will not necessarily have negative effects on access to and utilization of family planning, maternal health and child health services. Introducing user fees can even be associated with some positive effects on the utilization and the quality of ANC services. Fifth, any potential negative effects of introducing user fees in low- and middle-income settings on the utilization of healthcare services can be mitigated by officially exempting the poor from any fees at the point of service. More importantly, this exemption should be necessarily known to the population. Sixth, combining quality improvement interventions with user fees will not necessarily add to the few positive effects obtained when user fees are introduced as a stand-alone policy.

Seventh, provider incentives should be introduced carefully in low- and middle-income countries as negative effects are observed when these incentives are discontinued. To avoid these negative effects, adequate revenue streams should be generated to support the financial sustainability of the PBF schemes. In addition, consideration of alternative health sector interventions is initially required before introducing any provider incentives in low-resource settings, especially in light of the weak evidence on the maintenance of performance levels after PBF schemes are discontinued.

Finally, the lack of effects of different types and combinations of health sector interventions under Egypt's HSRP on modern contraceptive prevalence necessitates the investigation of the effectiveness of outreach activities in this regard.

7.4 LIMITATIONS OF THE STUDY

We realize that the study of this thesis has some limitations. These limitations are mainly related to the data used. First, the lack of random assignment to treatment and the lack of randomization from accreditation to contracting could affect the validity of our inferences. The best option to control for the HSRP targeting is to include the socio-economic indicators initially used for the targeting whenever applicable. In addition, a set of facility-level characteristics are included as controls whenever applicable to account for possible differences between treated and control facilities prior to being accredited, being contracted or shifting from accreditation to contracting.

Second, the Egypt DHS does not allow us to track the same women and children over time. Therefore, we collapse data from each DHS wave at the facility level and investigate the effects of interventions under the HSRP at the facility rather than the individual level.

Third, a limitation of spatially linking women to their nearest mapped health facilities is that we do not account for the fact that some women could seek health care from alternative sources apart from their nearest facilities. However, the phenomenon of bypassing is unlikely in our context as for a woman to use a public PHC facility in Egypt, she is obliged by the Ministry of Health (MOH) to use only the facility in catchment.

Fourth, we limit our analyses on child health to the *prevalence* instead of *treatment* of common early childhood illnesses. The number of observations that the Egypt DHS avails on treatment is statistically insufficient to construct indicators of treatment at the facility level. Thus, indicators of prevalence of these illnesses are constructed instead. Whenever applicable, we also construct a comprehensive indicator of early childhood mortality: the under-five mortality rate.

7.5 DIRECTIONS FOR FUTURE RESEARCH

The results of this thesis encourage an enquiry into several directions. First, future research is needed to understand the contextual factors associated with the differential effects of health sector interventions on family planning and maternal and child health. Our results can be complemented by in-depth process evaluations to identify mechanisms via which an intervention can or cannot be effective.

Second, the lack of effects of different interventions under Egypt's HSRP on modern contraceptive prevalence suggests investigation into the effectiveness of outreach activities, especially on family planning outcomes.

Third, the results of this thesis can be further strengthened by investigating the spillover effects of different interventions under the HSRP at the district level. This investigation would enable us to acquire a more comprehensive understanding of the effects of these interventions.

Fourth, it is also important for future research to investigate the differential effects of interventions on population subgroups, most importantly, targeted subgroups such as high-risk populations.

Fifth, future research should also attempt to investigate the effects of different health sector interventions on inequalities in access to and utilization of healthcare services and health outcomes.

Sixth, an extension of the work in this thesis can involve investigating the cost-effectiveness of health sector interventions in low- and middle-income countries in general and in Egypt in particular. Local governments, donor agencies and funding organizations seek evidence on the cost-effectiveness of different interventions. The financial sustainability of different intervention schemes, especially those that are externally funded, should be investigated as well.

References

Akerlof, G., 1970. The market for ‘lemons:’ Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics*, 84(3), pp.488-500.

Akin, J.S. & Hutchinson, P., 1999. Health-care facility choice and the phenomenon of bypassing. *Health Policy and Planning*, 14(2), pp.135-151.

Al Tehewy, M., Bssiouni, S., Habil, I. & El Okda, S., 2009. Evaluation of accreditation program in non-governmental organizations' health units in Egypt: Short-term outcomes. *International Journal for Quality in Health Care*, 21(3), pp.183-189.

Alkhenizan, A. & Shaw, C., 2011. Impact of accreditation on the quality of healthcare services: A systematic review of the literature. *Ann Saudi Med.*, 31(4), pp.407-416.

Angrist, J.D. & Krueger, A.B., 1999. Empirical strategies in labor economics. In: O. Ashenfelter & D. Card, eds. *Handbook of Labor Economics*, 3(A), pp.1277-1366. Amsterdam, New York and Oxford: Elsevier Science and North-Holland.

Arellano, M., 1993. On the testing of correlated effects with panel data. *Journal of Econometrics*, 59(1), pp.87-97.

Arrow, K., 1963. Uncertainty and the welfare economics of medical care. *American Economic Review*, 53, pp.941-969.

Ashenfelier, O. & Card, D., 1985. Using the longitudinal structure of earnings to estimate the effect of training programs. *Review of Economics & Statistics*, 67(4), pp.648.

Audibert, M. & Mathonnat, J., 2000. Cost recovery in Mauritania: Initial lessons. *Health Policy and Planning*, 15(1), pp.66-75.

Bagwell, K. & Michael, H.R., 1991. High and declining prices signal product quality. *American Economic Review*, 81, pp.224-239.

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

Becker, G.S., 1964. *Human capital: A theoretical and empirical analysis with special reference to education*. New York: National Bureau of Economic Research and Columbia University Press.

Becker, G.S., 1967. Human capital and the personal distribution of income. *W.S. Woytinsky Lecture No. 1*. Ann Arbor: Institute of Public Administration, University of Michigan.

Benjamin, A.L., Sapak, P., & Purai, J.K., 2001. User charges and utilisation of obstetric services in the National Capital District Papua New Guinea. *Pacific Health Dialogue*, 8(1), pp.38-43.

Ben-Porath, Y., 1967. The production of human capital and the life cycle of earnings. *Journal of Political Economy*, 75(4), pp.352.

Benzer, J.K., Young, G.J., Burgess, J.F., Baker, E, Mohr, D.C., Charns, M.P. & Kaboli, P.J., 2013. Sustainability of quality improvement following removal of pay-for-performance incentives. *Journal of General Internal Medicine*, 29(1), pp.127-132.

Bertrand, M., Duflo, E. & Mullainathan, S., 2004. How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics*, 119(1), pp.249-275.

Blundell, R., Duncan, A. & Meghir, C., 1998. Estimating labor supply responses using tax reforms. *Econometrica*, 66(4), pp.827-861.

Boland, G.W.L., Halpern, E.F. & Gazelle, G.S., 2010. Radiologist report turnaround time: Impact of pay-for-performance measures. *American Journal of Roentgenology*, 195(3), pp.707-711.

Bonfrer, I., Van, D.P. & Van Doorslaer, E., 2014. The effects of performance incentives on the utilization and quality of maternal and child care in Burundi. *Social Science & Medicine*, 123, pp.96-104.

Bratt, J.H., Weaver, M.A., Foreit, J., De Vargas, T. & Janowitz, B., 2002. The impact of price changes on demand for family planning and reproductive health services in Ecuador. *Health Policy and Planning*, 17(3), pp.281-287.

Brubakk, K., Vist, G.E., Bukholm, G., Barach, P. & Tjomsland, O., 2015. A systematic review of hospital accreditation: The challenges of measuring complex intervention effects. *BMC Health Services Research*, 15(1), 280.

Bukonda, N., Tavrow, P., Abdallah, H., Hoffner, K. & Tembo, J., 2002. Implementing a national hospital accreditation program: The Zambian experience. *International Journal for Quality in Health Care*, 14(Supplement 1), pp.7-16.

Caliendo, M. & Kopeinig, S., 2008. Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*, 22(1), pp.31-72.

Card, D. & Krueger, A.B., 1994. Minimum wages and employment: A case study of the fast-food industry in New Jersey and Pennsylvania. *American Economic Review*, 84(4), pp.772-793.

Card, D., 1992. Do minimum wages reduce employment? A case study of California, 1987-89. *Industrial and Labor Relations Review*, 46(1), pp.38-54.

Castro, R., Clementi, G.L. & MacDonald, G., 2004. Investor protection, optimal incentives, and economic growth. *The Quarterly Journal of Economics*, 119(3), pp.1131-1175.

Chawla, M. & Ellis, R.P., 2000. The impact of financing and quality changes on health care demand in Niger. *Health Policy and Planning*, 15(1), pp.76-84.

Chen, C.S., Liu, T.C., Chen, B. & Lin, C.L., 2014. The failure of financial incentive? The seemingly inexorable rise of cesarean section. *Social Science & Medicine*, 101, pp.47-51.

Cleveland, E.C., Dahn, B.T., Lincoln, T.M., Safer, M., Podesta, M. & Bradley, E., 2011. Introducing health facility accreditation in Liberia. *Global Public Health*, 6(3), pp.271-282.

Cohen, J. & Dupas, P., 2010. Free distribution or cost-sharing? Evidence from a randomized malaria prevention experiment. *The Quarterly Journal of Economics*, 125(1), pp.1-45.

Das, A., Gopalan, S.S. & Chandramohan, D., 2016. Effect of pay for performance to improve quality of maternal and child care in low- and middle-income countries: A systematic review. *BMC Public Health*, 16(1).

Easterly, W., Levine, R. & Roodman, D., 2004. Aid, policies, and growth: Comment. *American Economic Review*, 94(3), pp.774-780.

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

El-Jardali, F., Jamal, D., Dimassi, H., Ammar, W. & Tchaghchaghian, V., 2008. The impact of hospital accreditation on quality of care: Perception of Lebanese nurses. *International Journal for Quality in Health Care*, 20(5), pp.363-371.

El-Zanaty, F. & Way, A., 2001. *Egypt Demographic and Health Survey 2000*. Cairo and New York: Ministry of Health and Population, National Population Council, and Macro International.

Ensor, T., Chhun, C., Kimsun, T., McPake, B. & Edoka, I., 2017. Impact of health financing policies in Cambodia: A 20 year experience. *Social Science & Medicine*, 177, pp.118-126.

Finkelstein, A., 2002. The effect of tax subsidies to employer-provided supplementary health insurance: Evidence from Canada. *Journal of Public Economics*, 84(3), pp.305-339.

Fiorentini, G., Bruni, M.L. & Ugolini, C., 2013. GPs and hospital expenditures. Should we keep expenditure containment programs alive? *Social Science & Medicine*, 82, pp.10-20.

Flodgren, G., Gonçalves-Bradley, D.C. & Pomey, M.P., 2016. External inspection of compliance with standards for improved healthcare outcomes. *The Cochrane Database of Systematic Reviews*, 12(3).

Gertler, P., Giovagnoli, P. & Martinez, S., 2014. *Rewarding provider performance to enable a healthy start to life: Evidence from Argentina's Plan Nacer*. Policy research working paper 6884. Washington, DC: World Bank.

Graham, W.J., Bell, J.S. & Bullough, C.H., 2001. Can skilled attendance at delivery reduce maternal mortality in developing countries? In: V. DeBrouwere & W. Van Lerberghe, eds. Safe motherhood strategies: A review of the evidence. *Studies in Health Services Organisation and Policy*, 17(1), pp.97-130. Antwerp: ITG Press.

Greene, W.H., 2008. *Econometric Analysis*. 6th ed. New Jersey: Prentice Hall.

Grossman, M., 1972. On the concept of health capital and the demand for health. *Journal of Political Economy*, 80(2), pp.223-255.

Grossman, M., 2000. The human capital model. In: A.J. Culyer & J.P. Newhouse, eds. *Handbook of Health Economics*, 1, pp.347-408. Amsterdam, New York and Oxford: Elsevier Science and North-Holland.

Grun, R. & Ayala, J., 2006. *Impact evaluation of the Egyptian Health Sector Reform Project: Pilot phase*. Health, nutrition and population discussion paper 54512. Washington, DC: World Bank.

Hausman, J.A., 1978. Specification tests in econometrics. *Econometrica*, 46(6), pp.1251-1271.

Heckman, J.J., Ichimura, H. & Todd, P., 1998. Matching as an econometric evaluation estimator. *Review of Economic Studies*, 65(2), pp.261-294.

Hinchcliff, R., Greenfield, D., Moldovan, M., Westbrook, J.I., Pawsey, M., Mumford, V. & Braithwaite, J., 2012. Narrative synthesis of health service accreditation literature. *BMJ Quality & Safety*, 21(12), pp.979-991.

Hong, R., Mishra, V. & Fronczak, N., 2011. Impact of a quality improvement programme on family planning services in Egypt. *Eastern Mediterranean Health Journal*, 17(1), pp.4-10.

Hort, K., Djasri, H. & Utarini, A., 2013. *Regulating the quality of health care: Lessons from hospital accreditation in Australia and Indonesia*. Working paper 28. Melbourne: The Nossal Institute for Global Health.

Hysong, S.J., Khan, M.M. & Petersen, L.A., 2011. Passive monitoring versus active assessment of clinical performance. *Medical Care*, 49(10), pp.883-890.

Issifou, S. & Kremsner, P.G., 2004. Impact of increasing consultation fees on malaria in Africa. *The Middle European Journal of Medicine*, 116(9-10), pp.332-333.

Jacobs, B. & Price, N., 2004. The impact of the introduction of user fees at a district hospital in Cambodia. *Health Policy and Planning*, 19(5), pp.310-321.

Kipp, W., Kamugisha, J., Jacobs, P., Burhnam, G. & Rubaale, T., 2001. User fees, health staff incentives and service utilisation in Kabarole district Uganda. *Bulletin of the World Health Organization*, 79(11), pp.1032-1037.

Kontopantelis, E., Springate, D., Reeves, D., Ashcroft, D.M., Valderas, J.M. & Doran, T., 2014. Withdrawing performance indicators: Retrospective analysis of general practice performance under UK Quality and Outcomes Framework. *British Medical Journal*, 348, g330.

Kremer, M. & Miguel, E., 2007. The illusion of sustainability. *The Quarterly Journal of Economics*, 122(3), pp.1007-1065.

Lagarde, M. & Palmer, N., 2008. The impact of user fees on health service utilization in low-and middle-income countries: How strong is the evidence? *Bulletin of the World Health Organization*, 86(11), pp.839-848.

Lester, H., Schmittdiel, J., Selby, J., Fireman, B., Campbell, S., Lee, J., Whippy, A. & Madvig, P., 2010. The impact of removing financial incentives from clinical quality indicators: Longitudinal analysis of four Kaiser Permanente indicators. *British Medical Journal*, 340, c1898.

Matee, M. & Simon, E., 2000. Utilisation of dental services in Tanzania before and after the introduction of cost-sharing. *International Dental Journal*, 50, pp.69-72.

Mincer, J., 1974. *Schooling, experience, and earnings*. New York: National Bureau of Economic Research and Columbia University Press.

Mubyazi, G., Massaga, J., Kamugisha, M., Mubyazi, J.N., Magogo, G.C., Mdira, K.Y., Gesase, S. & Sukwa, T., 2006. User charges in public health facilities in Tanzania: effect on revenues, quality of services and people's health-seeking behaviour for malaria illnesses in Korogwe district. *Health Services Management Research*, 19(1), pp.23-35.

Nguyen, H.T.H., Hatt, L., Islam, M., Sloan, N.L., Chowdhury, J., Schmidt, J., Hossain, A. & Wang, H., 2012. Encouraging maternal health service utilization: An evaluation of the Bangladesh voucher program. *Social Science & Medicine*, 74(7), pp.989-996.

Pagan, A. & Ullah, A., 1999. *Nonparametric econometrics*. Cambridge: Cambridge University Press.

Peabody, J.W., Shimkhada, R., Quimbo, S., Solon, O., Javier, X. & McCulloch, C., 2014. The impact of performance incentives on child health outcomes: Results from a cluster randomized controlled trial in the Philippines. *Health Policy and Planning*, 29(5), pp.615-621.

Quimbo, S.A., Peabody, J.W., Shimkhada, R., Woo, K. & Solon, O., 2008. Should we have confidence if a physician is accredited? A study of the relative impacts of accreditation and insurance payments on quality of care in the Philippines. *Social Science & Medicine*, 67, pp.505-510.

Rajan, R.G. & Zingales, L., 1998. Financial dependence and growth. *American Economic Review*, 88(3), pp.559-586.

Richard, F., Ouedraogo, C., Compaore, J., Dubourg, D. & De Brouwere, V., 2007. Reducing financial barriers to emergency obstetric care: Experience of cost-sharing mechanism in a district hospital in Burkina Faso. *Tropical Medicine and International Health*, 12(8), pp.972-981.

Ridde, V., 2003. Fees-for-services, cost recovery, and equity in a district of Burkina Faso operating the Bamako Initiative. *Bulletin of the World Health Organization*, 81(7), pp.532-538

Riley, J.G., 2001. Silver signals: Twenty-five years of screening and signaling. *Journal of Economic Literature*, 39(2), pp.432-478.

Rosenbaum, P.R. & Rubin, D.B., 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), pp.41-55.

Salmon, J.W., Heavens, J., Lombard, C. & Tavrow, P., 2003. *The impact of accreditation on the quality of hospital care: KwaZulu-Natal Province Republic of South Africa*. Bethesda: United States Agency for International Development.

Schaffer, M.E. & Stillman, S., 2010. *xtoverid: Stata module to calculate tests of overidentifying restrictions after xtreg, xtivreg, xtivreg2 and xthtaylor*. [online] Available at: <http://ideas.repec.org/c/boc/bocode/s456779.html> [Accessed 12 November 2016].

Silverman, B.W., 1986. *Density estimation for statistics and data analysis*. London: Chapman & Hall.

Skiles, M.P., Curtis, S.L., Basinga, P., Angeles, G. & Thirumurthy, H., 2015. The effect of performance-based financing on illness, care-seeking and treatment among children: An impact evaluation in Rwanda. *BMC Health Services Research*, 15(1), 375.

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), pp.1518-1527.

Spilimbergo, A., 2009. Democracy and foreign education. *American Economic Review*, 99(1), pp.528-543.

United Nations, 2015. *Trends in contraceptive use worldwide 2015*. [pdf] New York: United Nations. Available at: <http://www.un.org/en/development/desa/population/publications/pdf/family/trendsContraceptiveUse2015Report.pdf> [Accessed 20 November 2016].

United Nations Children Fund, 2015. *Children in Egypt: A statistical digest*. [pdf] Cairo: United Nations Children Fund. Available at: <https://www.unicef.org/egypt/eg_UNICEF-2015-Children-in-Egypt-Statistical-Digest.pdf> [Accessed 22 November 2016].

Van De Poel, E., Flores, G., Ir, P. & O'Donnell, O., 2016. Impact of performance-based financing in a low-resource setting: A decade of experience in Cambodia. *Health Economics*, 25(6), pp.688-705.

Viswanathan, H., & Salmon, J.W., 2000. Accrediting organizations and quality improvement. *American Journal of Managed Care*, 6(10), pp.1117-1130.

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. *The Cochrane Database of Systematic Reviews*, 15(2).

Wooldridge, J.M., 2002. *Econometric analysis of cross section and panel data*. Massachusetts: MIT Press.

World Bank, 2004. *Egypt - Health sector reform and financing review*. [pdf] Washington, DC: World Bank. Available at: <http://documents.worldbank.org/curated/en/245241468770714134/pdf/411970EG0Healt1ver0P07631901PUBLIC1.pdf> [Accessed 15 April 2012].

World Bank, 2010. *Egypt - management and service quality in primary health care facilities in the Alexandria and Menoufia governorates*. [pdf] Washington, DC: World Bank. Available at: <http://documents.worldbank.org/curated/en/814731468021569255/pdf/693450ESW0P112000June025020100lerg.pdf > [Accessed 20 April 2012].

World Health Organization, 2015. *WHO Statement on caesarean section rates*. [pdf] Geneva: World Health Organization. Available at: <http://apps.who.int/iris/bitstream/10665/161442/1/WHO_RHR_15.02_eng.pdf > [Accessed 30 May 2016].

World Health Organization, 2016. *Global health observatory (GHO) data*. [online] Available at: <http://www.who.int/gho/en/> [Accessed 7 December 2016].