Making under-resourced health systems work for vulnerable women
and children: Antenatal care in Malawi

by

Martina Mchenga

Submitted in accordance with the requirements for the degree of

DOCTOR OF PHILOSOPHY (ECONOMICS)

at the

Faculty of Economic and Management Sciences

Stellenbosch University

Supervisors: Professor Ronelle Burger

Associate Professor Dieter von Fintel

March 2020

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my original work, that I am the authorship owner thereof (unless to the extent explicitly otherwise stated) and that I have not previously in its entirety or part submitted it for obtaining any qualification.

Signature: Martina Mchenga

Date: March 2020

Declaration by the candidate:

With regard to chapter two, the nature and scope of my contribution were as follows:

Nature of contribution	Extent of contribution (%)
I helped formulate the research question and develop the concept, conducted data analysis, wrote first and final draft and submitted the paper for publication at the BMC Health Services Research journal.	70%

The following co-authors have contributed to chapter two:

Name	Email address	Nature of contribution	Extent of contribution (%)
Ronelle Burger	_____	Research design and manuscript editing	15%
Dieter von Fintel	_____	Research design and manuscript editing	15%

Signature of candidate: Declaration with signature in possession of candidate and supervisor.

Date: 25 October 2019

Declaration by co-authors:

The undersigned hereby confirm that:

1. The declaration above accurately reflects the nature and extent of the contributions of the candidate and the co-authors to chapter two.

2. No other authors contributed to chapter two besides those specified above.

3. Potential conflicts of interest have been revealed to all interested parties and the necessary arrangements have been made to use the material in chapter two of this dissertation.

Signature	Institution affiliation	Date
Declaration with signature in possession of candidate and supervisor.	Stellenbosch University	25 October 2019
Declaration with signature in possession of candidate and supervisor.	Stellenbosch University	25 October 2019

Declaration by the candidate:

With regard to chapter three, the nature and scope of my contribution were as follows:

Nature of contribution	Extent of contribution (%)
I helped formulate the research question and develop the concept, conducted data analysis and detailed write-up of the analysis.	70%

The following co-authors have contributed to chapter three:

Name	Email address	Nature of contribution	Extent of contribution (%)
Ronelle Burger	_____	Research design and manuscript editing	15%
Dieter von Fintel	_____	Research design and manuscript editing	15%

Signature of candidate: Declaration with signature in possession of candidate and supervisor.

Date: 25 October 2019

Declaration by co-authors:

The undersigned hereby confirm that:

1. The declaration above accurately reflects the nature and extent of the contributions of the candidate and the co-authors to chapter three.

2. No other authors contributed to chapter three besides those specified above.

3. Potential conflicts of interest have been revealed to all interested parties and the necessary arrangements have been made to use the material in chapter three of this dissertation.

Signature	Institution affiliation	Date
Declaration with signature in possession of candidate and supervisor.	Stellenbosch University	25 October 2019
Declaration with signature in possession of candidate and supervisor.	Stellenbosch University	25 October 2019

Abstract

This thesis investigates the implementation barriers and adoption of maternal health care interventions in Malawi, a sub-Saharan African country with one of the highest maternal mortality rates and most poorly resourced health systems in the region. Although antenatal care has universal components that apply to all pregnant women, the guidelines are designed to be adaptable so that countries with different health system structures and burdens of disease can implement them according to their context and the needs of their population (Benova et al., 2018). However, a lack of empirical evidence, means it is difficult to know and assess whether the existing models of care are being successfully implemented and, furthermore, whether they achieve the intended – or even unintended – objectives and provide solutions for the future, especially in settings where resources are limited. This thesis attempts to address this gap.

I focused on two major aspects of antenatal care (ANC): the supply aspect (quality of care) and the demand aspect (utilisation of care), to analyse three broad objectives. In the first objective, I examine the impact of the 2001 Focused Antenatal Care (FANC) model on quality of care and utilisation of ANC services in Malawi. In the second objective, I estimate the optimal number of visits that are effective in improving birth outcomes in low-resourced settings. In the third objective, I compare women's self-reports on the quality of ANC received to the direct observation of facilities to understand how to counter biases and mismeasurements that can impede accurate local tracking of the quality of services provided.

An examination of the impact of the 2001 World Health Organization FANC model on the utilisation, early access and quality of care in Malawi, using three comparable demographic and health datasets, and the interrupted time series analysis, confirms that, when it comes to policy, one size does not always fit all. The findings reveal that, although FANC only recommends a minimum number of four visits and is therefore potentially cost effective, in Malawi the model did not translate into improved quality of care and was found to be associated with the unintended consequence of increased underutilisation of ANC. It is therefore questionable whether revising the minimum number of visits from four to eight, as recommended by the 2016 WHO guidelines would be effective in low-resourced settings.

Villar et al. (2002) noted that most ANC policies in low and medium-income countries (LMICs) are adopted without thorough scientific evaluation and that there is a lack of empirical evidence on the average number of visits likely to produce the most benefit in improving maternal and child health outcomes. Therefore, this thesis extends the analysis from objective one and estimates the number of visits that would be effective to improve birthweight in Malawi. Using nationally representative Malawi Demographic and Health Survey data, I apply instrumental variable models together with highly flexible non-linear spline specifications and Wald tests to estimate breaks in the relationship between the number of ANC visits and the probability of low birthweight. Results suggest that only three visits are required to reduce the probability of low birthweight to the same extent as more visits would. This implies that low-income health systems are likely to perform just as well if fewer routine visits are conducted with more attention to quality, and reserving additional ANC visits to women who critically need them.

The analysis in the first two objectives takes advantage of the publicly available nationally representative household Malawi Demographic and Health Surveys, which are based on women's self-reports on the services provided. However, the reliability of this data depends on a number of factors: the client's ability to recall with accuracy; the client's access to

information and knowledge of the content of care; an understanding of the questions being asked and the ability to link them to what the provider was doing, among others. These highlighted limitations may lead to an upward or downward bias in the quality of care measurement, limiting the utility of survey results for programme improvement. Given that most LMICs rely on household and client exit surveys to obtain estimates of healthcare quality, accurate information on ANC quality is important.

In this part of the thesis, therefore, I assess the extent to which women's self-reports on the quality of ANC is congruent with facility-observed estimates by testing the sensitivity, specificity and receiver-operating curves (ROCs) of ANC quality components. The results suggest that women overestimate the quality of care, mainly due to a lack of knowledge about complicated items of ANC and an overall understanding of the questions beings asked in the surveys. For example, complicated quality components asking whether the provider had counselled the client on the side effects of iron and pregnancy-related complications had lower reporting accuracy than objective indicators asking whether the provider had prescribed medication for malaria prevention, and iron/folic tablets. The main recommendations in this regard are that, in measuring the quality of care, it is important to compare women's self-reports to facility data to get accurate quality estimates; and that the emphasis needs to be on women to place them at the forefront of policy change by educating them on what to expect during healthcare consultations.

The overall findings suggest that public policy has an important role to play when it comes to maternal preventive healthcare. Demand-side policy tools such as increased access to basic information on the importance of accessing ANC services and specific components to expect during a consultation can only be successful if the supply side is adequate and effective.

Opsomming

Hierdie tesis ondersoek hindernisse tot die implementering van voorgeboortelike sorg en die benutting daarvan deur swanger vroue in Malawi. Dié Afrikaland suid van die Sahara het een van die hoogste sterftesyfers van swanger vroue in die streek, sowel as 'n disfunksionele gesondheidstelsel met onvoldoende hulpbronne. Hoewel roetine voorgeboortelike sorg algemene komponente insluit wat op elke swanger vrou van toepassing is, is die amptelike aanbevelings so geformuleer dat verskillende lande – met verskillende ladings van siekte, sosio-ekonomiese kontekste en gesondheidstrukture – dit kan aanpas en volgens hulle konteks en bevolkings se behoeftes kan implementeer. Daar is egter 'n gebrek aan empiriese bewyse wat dit moeilik maak om te bepaal of die bestaande modelle wat vir voorgeboortelike sorg gebruik word, wel suksesvol geïmplementeer word; of dit die voorgenome – of dalk onbeplande – doelwitte bereik, en of dit oplossings vir die toekoms skep, veral in lande met beperkte hulpbronne. Hierdie tesis poog om dié gaping in die navorsing te vul.

Die studie ontleed drie breë doelwitte deur te fokus op twee hoofaspekte van voorgeboortelike sorg: aanbod (kwaliteit van sorg) en vraag (benutting van sorg). Die eerste doelwit ondersoek die impak van die Wêreld Gesondheidsorganisasie (WGO) se 2001 Gefokusde Voorgeboortelike Sorg- (*Focused Antenatal Care*) (FANC-) model op die gehalte en gebruik van voorgeboortelike dienste in Malawi. Die tweede doelwit beraam die optimale aantal besoeke wat nodig is om geboorte-uitkomste in lande met lae hulpbronne doeltreffend te verbeter. Wat die derde doelwit betref, word vroue se selfverslae met die direkte waarneming van fasiliteite vergelyk ten einde te verstaan hoe om vooroordele en mismetings, wat akkurate plaaslike monitering van die gehalte van dienste kan belemmer, die hoof te bied.

'n Ondersoek is gedoen oor die impak van die FANC-model op die benutting van vroeë toegang tot en die gehalte van voorgeboortelike sorg in Malawi. Die ondersoek het van drie vergelykbare demografiese en gesondheidsorgdatastelle gebruik gemaak en bevind dat een beleid nie vir alle situasies geskik kan wees nie. Hoewel FANC 'n minimum van vier besoeke aanbeveel, wat voorgeboortelike sorg potensieel koste-effektief maak, is daar bevind dat die model nie die gehalte van voorgeboortelike sorg in Malawi verbeter het nie. Inteendeel, FANC het 'n onbeplande uitkoms gehad, naamlik verhoogde onderbenutting van voorgeboortelike sorg. Die vraag ontstaan dus of die hersiening van die minimum aantal besoeke deur die 2016 WGO-riglyne, van vier na agt, in laehulpbron-situasies doeltreffend sal wees.

Villar et al. (2002) het opgemerk dat, in lande met 'n lae en medium inkomste (LMI), beleide oor voorgeboortelike sorg meestal sonder deeglike wetenskaplike evaluering geïmplementeer word. Daar is ook nie genoeg empiriese bewyse om aan te dui gemiddeld hoeveel besoeke gesondheidsuitkomste vir moeders en kinders sal verbeter nie. Teen dié agtergrond poog hierdie tesis om 'n beraming te maak van die optimale aantal besoeke wat die geboortemassa van babas in Malawi sal verhoog. Met behulp van nasionaal verteenwoordigende data uit Malawi se Demografiese en Gesondheidsopname, is instrumentele veranderlike modelle tesame met hoogs buigsame nie-liniêre latfunksie-spesifikasies en Wald-toetse aangewend om die breekplekke in die verhouding tussen die aantal voorgeboortelike besoeke en die waarskynlikheid van lae geboortemassa te beraam. Resultate dui daarop dat slegs drie besoeke nodig is om die waarskynlikheid van lae geboortemassa te verminder, in dieselfde mate as wat meer besoeke dit sou verminder. Minder roetine besoeke van beter gehalte sal dus moontlik net so goed werk in LMI-lande as meer besoeke (maar van laer gehalte). Bykomende

voorgeboortelike besoeke behoort dan eerder benut te word vir vroue wat werklik ernstige probleme het.

Die Malawi Demografiese en Gesondheidsopnames wat vir die eerste twee doelwitte gebruik word, is gebaseer op vroue se selfverslae oor die voorgeboortelike dienste wat hulle ontvang het. Die betroubaarheid van die data berus egter op 'n aantal faktore, onder andere die kliënt se vermoë om presies te onthou; die kliënt se toegang tot inligting en kennis van wat voorgeboortelike sorg behels; 'n begrip van die vrae wat gevra is en die vermoë om die vrae te koppel aan dit wat die sorgverskaffer in werklikheid gedoen het. Hierdie beperkings kan lei tot 'n positiewe of negatiewe vooroordeel wanneer die gehalte van sorg gemeet word. Dit kan die gebruik van opname-uitslae onbetroubaar maak as dit vir programverbetering gebruik word. Gegewe dat die meeste LMI-lande op huishouding- en kliëntuitgang-opnames staatmaak om die gehalte van gesondheidsorg te bepaal, is akkurate inligting oor die gehalte van voorgeboortelike sorg noodsaaklik.

In hierdie deel van die tesis is daar dus bepaal in hoe 'n mate vroue se selfverslae oor die gehalte van voorgeboortelike sorg ooreenstem met fasiliteitwaarnemings van dienste wat werklik gelewer is. Daar is getoets vir sensitiwiteit, spesifisiteit en ontvanger-keuringskurwes (*receiver-operating curves*) (ROCs) ten opsigte van die gehaltekomponente van voorgeboortelike sorg. Die resultate dui daarop dat vroue die gehalte van sorg oorskat, hoofsaaklik weens 'n gebrek aan kennis oor komplekse aspekte van voorgeboortelike sorg en 'n swak algehele begrip van die vrae wat in die opnames gevra is. Byvoorbeeld, vrae oor komplekse gehaltekomponente soos of die verskaffer die kliënt ingelig het oor die newe-effekte van yster en oor swangerskapverwante komplikasies het laer verslaggewingsakkuraatheid getoon as objektiewe aanwysers wat gevra het of die verskaffer medikasie vir malariavoorkoming en yster- of foliensuurtablette voorgeskryf het. Wat die meting van die gehalte van sorg betref, word daar dus aanbeveel dat vroue se selfverslae met fasiliteitsdata vergelyk moet word ten einde 'n akkurate beraming van gehalte te maak. Vroue moet ook die vertrekpunt vir beleidsverandering wees deur hulle op te voed oor wat om tydens gesondheidsorgkonsultasies te verwag.

Die algehele bevindinge van hierdie studie dui daarop dat openbare beleid 'n belangrike rol in voorkomende gesondheidsorg vir swanger vroue kan speel het. Aan die vraagkant kan beleidshulpmiddels, soos beter toegang tot basiese inligting oor voorgeboortelike dienste en oor die spesifieke komponente wat tydens 'n konsultasie verwag kan word, slegs suksesvol wees indien die aanbodkant se lewering voldoende en doeltreffend is.

Acknowledgements

I express profound gratitude to my supervisors, Professor Ronelle Burger and Associate Professor Dieter von Fintel, for their guidance and support throughout my research. Their valuable input and timely assistance on this study helped to shape this work into what it is now.

I also gratefully acknowledge the financial support towards my studies and thesis support by the Stellenbosch University Graduate School of Economics and Management Sciences (GEM), the Canon Collins Scholarship for Postgraduate Studies in South Africa and the Desmund Tutu. Being part of the GEM programme and Research on Socio-Economic Policy (ReSEP) has equipped me with research skills, experience and exposure. I am humbled to have met and learnt from my fellow PhD students and researchers during the weekly seminars.

I am grateful for the comments and suggestions I got from the GEM and ReSEP weekly research seminars, the early careers session organised by the 2019 biennial International Health Economics Association (IHEA) Conference, the fifth Global Symposium on Health Systems Research (HSR2018), and reviewers at the *BMC Health Services Research* journal who reviewed my PhD article in chapter two before it got published.

Special thanks should go to my parents and siblings for their constant support, love and belief in me throughout my studies. To my partner Upile: thank you for being a pillar of encouragement and physically being there for me, especially at the very end of my studies; and to my SIF family in Stellenbosch and friends back home, thanks for all your love, support and encouragement.

To my daughter, Ndamo, thank you for showing me that I could do both a career and be your Mommy. This work is for you and I love you beyond words.

Above all, I thank the Almighty God, for seeing me through and fulfilling His promises upon my life throughout my studies. Thus far, you have brought me!

CONTENTS

List of figures

List of tables

Abbreviations and acronyms

ANC	Antenatal care
ARV	Antiretroviral
CHAM	Christian Health Association of Malawi
CI	Confidence interval
DHS	Demographic and Health Survey
EHP	Essential healthcare package
FANC	Focused Antenatal Care
FPE	Free primary education
GDP	Gross domestic product
GEM	Graduate School of Economics and Management
IHEA	International Health Economics Association
ITSA	Interrupted time series analysis
LBW	Low birthweight
LMIC	Low and medium-income countries
MCA	Multiple correspondence analysis
MDG	Millennium Development Goal
MICS	Multiple Indicator Cluster Survey
MMR	Maternal mortality rate
MoH	Ministry of Health
OLS	Ordinary least squares
PMTCT	Prevention of mother-to-child transmission of HIV
ReSEP	Research on Socio-Economic Policy
ROC	Receiver-operating curve
SDG	Sustainable Development Goals
SPA	Service Provision Assessments
TBA	Traditional birth attendant
WHO	World Health Organization

CHAPTER ONE

Introduction

1.1   Overview

In the past few decades, significant progress in the reduction of child and maternal mortality has been made worldwide. Globally, in the past 25 years, maternal mortality rate (MMR) dropped by almost 44% (WHO, 2015a) and the under-five mortality by 56% (UN IGME, 2017). Notwithstanding this progress, the survival of mothers and children remains an urgent concern, especially in poor resourced settings. In 2015, approximately 303 000 women and adolescent girls lost their lives to complications during pregnancy and childbirth (WHO, 2015a). Similarly, there were 5.6 million under-five deaths, 2.6 million (46%) of them in the first 28 days of life. About 99% of maternal deaths (WHO, 2015a) and 80% of under-five deaths (UN IGME, 2017) occur in countries with constrained resources. The good news is that, most of these adverse outcomes can be prevented with good-quality care, including antenatal care (ANC) (Benova et al., 2018).

Empirical evidence has shown that the effective use of antenatal care is among the preventive interventions to improve health outcomes, reduce maternal and neonatal mortality, reduce postpartum anaemia and ensure appropriate birthweight (WHO, 2009; WHO, 2015a; Adekanle & Isawumi, 2008; Khatun & Rahman, 2008). According to the 2016 WHO ANC guidelines, antenatal care offers an opportunity for the health provider to monitor and ensure the well-being of both the mother and the foetus as well as detect any pregnancy-related complications and take the necessary precautions (WHO, 2016). Furthermore, it also provides an opportunity to prepare the mother for birth and overall promote healthy behaviours of the mother (WHO, 2016). Studies by Testa et al. (2002) and Prual et al. (2002) estimated that antenatal care alone could reduce maternal mortality by 20%, provided that the care was of good quality and mothers made regular antenatal visits.

The bad news, however, is that the implementation of antenatal care interventions is ineffective and the demand for antenatal care, just like other preventive interventions, is often low in LMICs (Dupas, 2011). For example, in 2001, the World Health Organization (WHO) began promoting a new model of antenatal care, which emphasised quality of care and reduced the

minimum number of visits to four for women with uncomplicated pregnancies. More than a decade after the policy change, only 52% of women had received at least four antenatal care visits in sub-Saharan countries (UNICEF, 2018). Understanding the barriers to the implementation and adoption of maternal healthcare interventions is a critical issue in achieving the United Nations' maternal and child health-related Sustainable Development Goals (SDGs) and in development economics at large. This thesis provides microeconomic evidence on this issue.

Specifically, the thesis examines two major aspects of antenatal care in Malawi: the supply aspect (quality of care) and the demand aspect (utilisation of antenatal care), to analyse three broad objectives. In the first objective, I examine the impact of the 2001 Focused Antenatal Care (FANC) model on the quality and utilisation of antenatal care services in Malawi. FANC was adopted in Malawi in 2003, replacing the traditional antenatal care model, which included numerous visits (7–16 visits) (WHO, 2002a). Evidence shows that the four-visit model had substantial public health implications, especially in low-income countries where healthcare resources are inadequate (Villar et al., 2001). Moreover, in the 2015 Cochrane review, Dowswell et al. (2015) argued that the FANC model reduces the costs for women, by reducing travel times to the clinic and loss of working hours. However, the impact of FANC on the utilisation and quality of antenatal care services in low-resourced settings has been inconclusive.

In the second objective, I estimate the optimal number of antenatal care visits that are effective in improving birth outcomes in low-resourced settings. Finally, in the last objective, I explore the information asymmetry between providers and clients by comparing women's self-reports on the quality of antenatal care received with direct observations of facilities to understand how to counter biases and mismeasurements that can impede accurate local tracking of the quality of services provided.

I specifically focus on this area to contribute to and engage with the debate on the appropriate models of antenatal care in low-resourced settings and appropriate data sources for measuring progress towards quality universal health access. For the Sustainable Development Goals to be achieved in the coming decade, the global community needs to focus on, invest in and promote understanding of how health policies and systems can be strengthened.

1.2   Background and study context

1.2.1   Malawian healthcare system and healthcare delivery

According to the 2014 Malawi Ministry of Health report, the Malawian healthcare system has a three-tier healthcare delivery system (MoH, 2014b). These tiers are linked through an elaborate referral system (WHO, 2008a). The first tier, primary healthcare, consists of smaller level facilities such as health posts, dispensaries, maternity units, health centres, and community and rural hospitals (MoH, 2014b). The second tier constitutes district hospitals and provides specialised services to patients referred from the primary healthcare level through outpatient and inpatient services and community healthcare services (MoH, 2014b). The third tier, tertiary healthcare provides the highest level of healthcare. It consists of highly specialised services and is provided by central and other specialist hospitals. In practice, however, about 70% of the tertiary healthcare services are either primary or secondary services because there is not an effective gate-keeping system (MoH, 2011).

The Ministry of Health (MoH) is the main provider of healthcare services in Malawi, owning nearly half (48%) of the facilities (MoH, 2014a). This is followed by healthcare facilities run by private for-profit institutions (Figure 1.1).
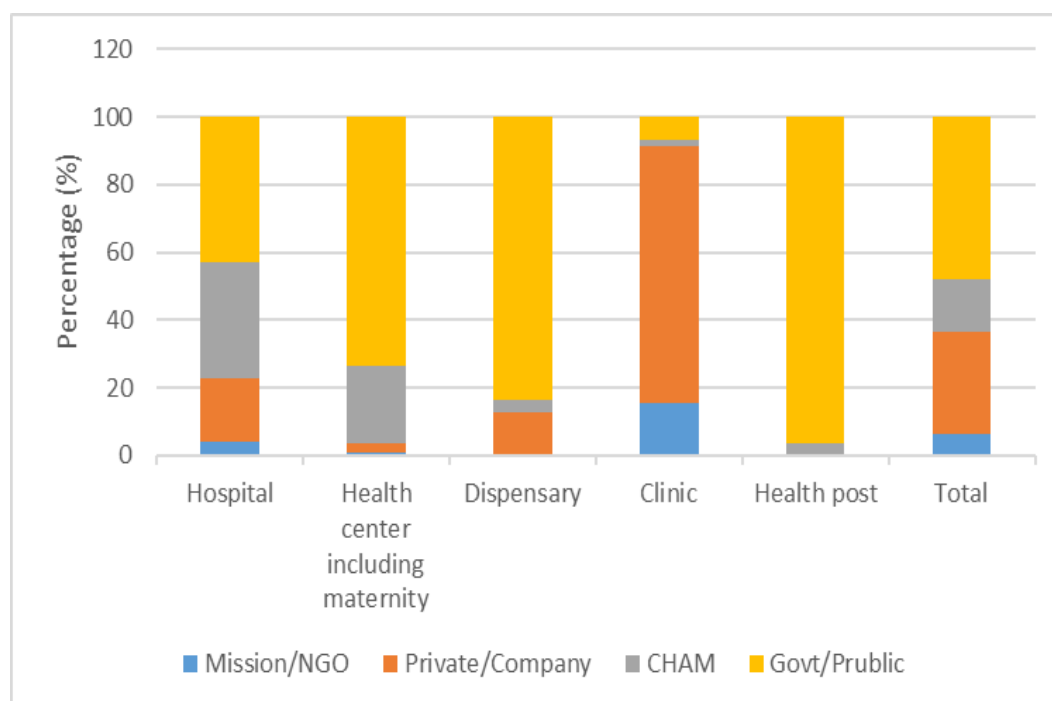
Figure 1.1: Health facilities by managing authority
Source: Own calculations from the 2013/2014 SPA

3

One of the key priorities of the Malawi government is to have the highest possible level of health and quality of life for all its citizens (GoM, 2017). To achieve this priority, primary healthcare services are provided free in all government facilities through the essential healthcare package (EHP). However, the Malawian healthcare system faces many challenges to effectively provide primary healthcare services. Economically, the country's gross domestic product (GDP) growth rate declined from 9.5% in 2010 to 5.8% in 2015 (World Bank, 2015). This decline resulted in budget cuts, meaning that the health sector no longer received sufficient support from the government. In 2017, health expenditure per capita was $39.20 in Malawi. This is more that 50% less than the WHO recommendation that governments need to spend at least $86 per person to provide people with the essential healthcare package (UNICEF, 2018). Furthermore, in recent years, as shown in Figure 1.2, the Malawi government has failed to meet the 15% of total budget threshold set by the African Union's Abuja Declaration in 2001 to spend on health.
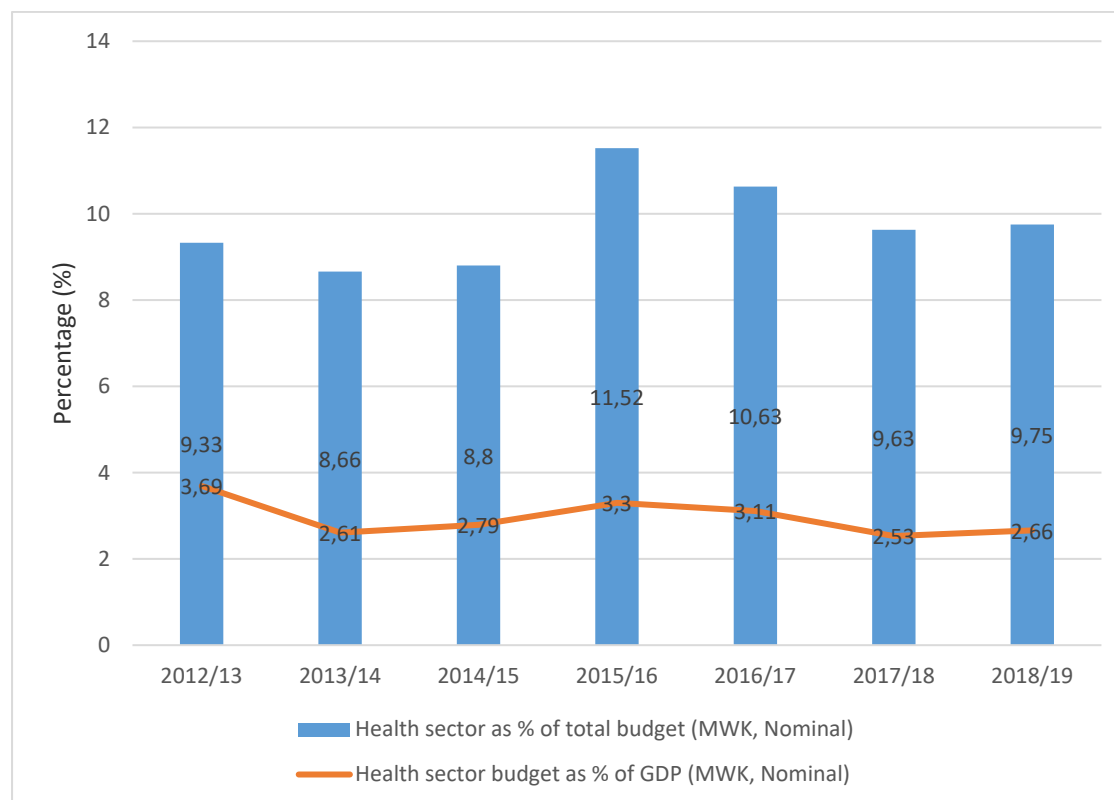
Figure 1.2: Health sector spending as a percentage of total budget and of total GDP
Source: Own calculations based on UNICEF report (2018)

Moreover, despite high poverty levels, the out-of-pocket expenditure of households for health stands at 24% of total household income, which is one of the highest in Africa (GoM, 2017). This contradicts the universal health coverage policy on health financing, which recommends

that countries reduce the household financial burden of accessing care. Besides financial resources, the Malawian healthcare system also faces challenges in terms of huge shortages in health worker resources (Muula, 2006). In 2017, Malawi had a population of over 17 million (NSO, 2018). In the same year, there were only 284 medical officers and 1 159 clinical officers nationwide (GoM, 2017). In terms of nursing professionals, the country reported 1 098 nursing officers and only 3 475 midwife technicians (GoM, 2017). This amounts to about 0.283 nurses and midwives for every 1 000 people in the country, whereas the WHO's standard is 2.5 skilled health workers per 1 000 (Global Fund, 2019). While over 80% of the population resides in rural areas, only 29% of nursing professionals and 40% of associate nursing professionals are allocated in rural areas (GoM, 2017).

Suffice to say that Malawi's health system faces absolute and relative inadequacy of financing to fund the free primary healthcare services, undermining the government's ability to provide universal health coverage for all. This aggravates the undesirably low level of maternal and child outcomes, especially among the vulnerable women and children in Malawi. In addition, the capacity deficiency experienced in the healthcare system creates a very challenging environment for successful policy implementation. It can affect the effectiveness of the policy itself in achieving the intended objectives.

1.2.2   Maternal and child healthcare in Malawi

Over the years, Malawi has made significant improvements in reducing maternal and neonatal mortality rates (Figure 1.3)[1], nevertheless, the country failed to achieve the 2015 Millennium Development Goal (MDG) of reducing the maternal mortality rate (MMR) by 75% and the neonatal mortality to 12 per 1 000 live births. With the advent of the Sustainable Development Goals (SDGs), it remains to be seen whether Malawi will be able to achieve the ambitious SDG targets: reducing the MMR to below 70 per 100 000 live births and neonatal mortality to 12 per 1 000 by 2030. Regarding the MMR, the WHO argues that to achieve the SDG target of 70 per 100 000 live births by 2030, countries will require to reduce their MMR by at least 7.5% every year between 2016 and 2030 (WHO, 2015a). Malawi managed to reduce its MMR by 1.6% per year between 1990 and 2015.

---

[1] Given data limitations, I was however, not able to show whether this decline in MMR was due to the focused antenatal care (FANC) or not.

In Malawi, the majority of the maternal and neonatal deaths are attributed to haemorrhage, low birthweight, sepsis, eclampsia and premature delivery (WHO, 2005), most of which can be prevented through the provision of antenatal care and skilled delivery care. In order to improve maternal and child health outcomes, Malawi launched the Presidential Initiative on Maternal Health and Safe Motherhood in 1998 (Malata, 2016). One of the main pillars of this initiative is antenatal care.
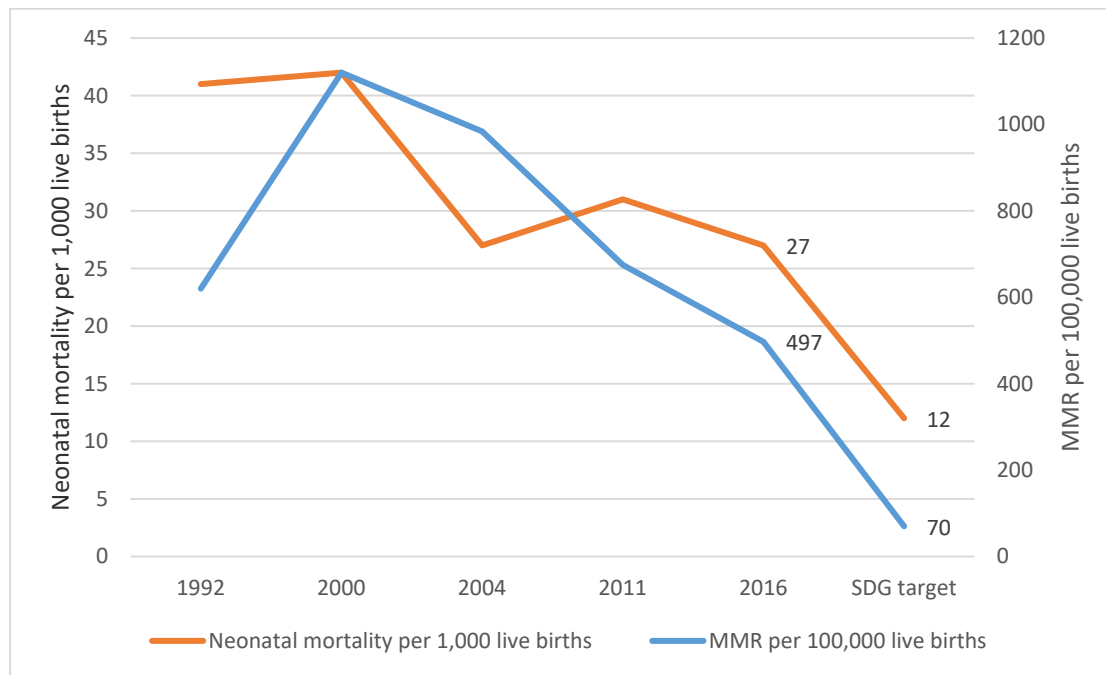
Figure 1.3: Trends in maternal and neonatal mortality rates
Source: Own calculations using the Malawi DHS 1992, 2000, 2004, 2010/2011, 2016

Over the past decades, the WHO has proposed various antenatal care models. High-income countries use a traditional model of antenatal care that involves more frequent and a large number of antenatal care visits – approximately 7–16. Pregnant women start antenatal care as early as possible in these countries, with monthly visits up to 28 weeks and weekly visits up to 36 weeks, until delivery (Say & Raine, 2007). Furthermore, pregnant women have access to adequate antenatal care, which includes regular tests and ultrasound assessments.

However, owing to the high cost, the traditional model of antenatal care did not work in low-resourced settings (Gajate-Garrido, 2013). In 2001, the WHO therefore began promoting a new model of antenatal care for LMICs called Focused Antenatal Care (FANC). Unlike the traditional model, FANC recommends a minimum of four visits for women with an uncomplicated pregnancy. It focuses on goal-oriented and targeted care to better detect and manage complications during pregnancy (WHO, 2002a). According to FANC, the first visit

6

should occur by week 16; the second at weeks 24–28; the third at week 32; and the final one by weeks 36–38 (WHO, 2002a). Malawi adopted FANC in 2003 and this is the antenatal care model currently in use (MoH, 2014b).

Study results on the effectiveness of FANC have been inconclusive. While other researchers reported that FANC was safe, sustainable, comprehensive and a cost-effective antenatal care model (Villar et al., 2001; Birungi & Onyango-Ouma, 2006; Nyarko et al., 2006), in 2015, a systematic review found that the reduced model of visits was associated with increased perinatal deaths compared to models that recommended at least eight visits (Dowswell et al., 2015). This prompted the WHO to revise the antenatal care policy in 2016, recommending an increase in the number of visits to a minimum of eight contacts (WHO, 2016). In the revised model, the first visit is recommended to take place in the first trimester by 12 weeks of gestation. During the third trimester, which is the time of highest risk for pre-eclampsia and eclampsia, five visits should be scheduled as follows; at 30, 34, 36, 38 and 40 weeks. Furthermore, the new antenatal care guidelines include 49 extensive recommendations that cover a wide range of interventions (WHO, 2016).

In 2016, *The Lancet* initiated the debate on whether it is advisable for LMICs to mobilise the resources required to double the minimum number of visits or contacts from four to eight (Weeks & Temmerman, 2016). The problem with most preventive healthcare interventions such as antenatal care is that households in low-income countries tend to underinvest in them (Dupas, 2011). For example, in Malawi, despite the fact that antenatal care is provided free in all public facilities in the country, giving antenatal care coverage of over 90% in both rural and urban areas, only 50% of pregnant women made at least four antenatal care visits in 2016 (Table 1.1). These statistics are disappointingly low. In this context, it is puzzling that women are not making the adequate number of antenatal care visits as required by FANC.

To explain the puzzle, Dupas (2011) argues that, in some cases where households underinvest in preventive measures, the problem may be the supply. She argues that the delivery of preventive care, which is mostly through the public sector, is often poor in low-resourced settings (Dupas, 2011). For example, studies assessing the implementation of FANC in LMICs found that most LMICs failed to implement the model effectively, which means that FANC did not improve the quality of antenatal care services (Mchenga et al., 2019; Chege, 2005; Nyarko et al. 2006). Some of the reasons for the failure were inadequate equipment, supplies,

infrastructure and training (Lungu et al., 2011; Birungi & Onyango-Ouma, 2006; Nyarko et al., 2006; Chege, 2005).

Table 1.1: Utilisation of ANC services by type of residence

	Rural	**Urban**	**National**
2000			
ANC by skilled professional (nurse/doctor)	**90.6**	**97.3**	**91.0**
No access (did not initiate ANC)	5.0	1.5	4.6
First visit by 4 months	6.4	7.5	6.5
4+ ANC visits	54.1	68.3	56.0
2004/2005			
ANC by skilled professional (nurse/doctor)	**91.2**	**97.6**	**93.0**
No access (did not initiate ANC)	5.1	1.9	4.6
First visit by 4 months	7.4	9.4	7.7
4+ ANC visits	55.1	65.2	57.1
2010			
ANC by skilled professional (nurse/doctor)	**94.0**	**96.0**	**95.0**
No access (did not initiate ANC)	1.6	1.4	1.6
First visit by 4 months	12.4	12.6	12.4
4+ ANC visits	44.9	48.6	45.5
2015/2016			
ANC by skilled professional (Nurse/doctor)	**94.0**	**97.0**	**95.0**
No access (did not initiate ANC)	2.0	1.0	1.8
First visit by 4 months	23.7	25.5	24.0
4+ ANC visits	49.2	58.9	50.6

Source: Malawi DHS 2000, 2004, 2010, 2015/2016

Providers in LMICs have also raised concerns about the difficulty of incorporating all the FANC protocols into relatively short appointments (Maternal Health Task Force, 2014). Given the challenges faced by LMICs in implementing FANC and achieving the less ambitious target of a minimum of four visits, it is questionable whether the 2016 antenatal care guidelines would be effective in settings where resources are limited. Villar et al. (2001) argued that most antenatal care policies in LMICs are adopted without thorough scientific evaluation. Thus, there is a need for empirical evidence on the average number of visits that are likely to produce the most benefit in improving maternal and child health outcomes in settings where resources are limited.

Another reason why households in LMICs underinvest in preventive healthcare as argued by Dupas (2011) is the lack of information on how to prevent illness and the cost-effectiveness of preventive behaviour. Supporting this argument, Arrow (1983), noted that in most LMICs,

market imperfections are common, in particular information asymmetry between the health provider and the client. Both Dupas (2011) and Arrow (1983) argued that, in low resourced settings, information asymmetry makes access to the information by the client difficult. The main reasons are the low penetration of media communication on public health topics, low levels of education and poor access to healthcare services. Empirical research has shown that access to information can have a positive impact on health behaviour and that media campaigns about specific prevention practices could make a difference in household behaviour and increase the demand for health services (Dupas, 2011). For example, in LMICs, studies have reported that increased access to information through mass media in turn increased the utilisation of antenatal care services (Zamawe et al., 2016; Archarya, et al., 2015; Edward, 2011; Kulkarni & Nimbalkar, 2008).

On the other hand, if women do not have adequate information, they may not know the importance of antenatal care and what to expect at the antenatal care visits, and may therefore decide to forgo antenatal care altogether. Moreover, a lack of adequate information on specific interventions a woman should expect during an antenatal care visit can potentially affect the reporting accuracy of the quality of care received. In LMICs, inaccurate information on the demand for services and the quality of care from clients has negative policy implications because estimates of the demand for health services and the quality of care are mainly calculated from household and client exit surveys, which are based on self-reported data (Blanc et al., 2016). If information is reported with errors, it limits the utility of the surveys in programme improvement (Lindelow, 2003).

1.3   Research questions and methods used

This thesis investigates the implementation barriers and adoption of maternal health care interventions in Malawi, a sub-Saharan African country with one of the highest maternal mortality rates and most poorly resourced health systems in the region. Although antenatal care has universal components that apply to every pregnant woman, the guidelines are designed to be adaptable so that countries with different health system structures and burdens of disease can implement them according to their context and the needs of their population (Benova et al., 2018). A lack of empirical evidence in low-resourced settings, however, means that it is difficult to know and assess whether the existing models of care are successfully implemented

and achieve the intended or unintended objectives and provide solutions for the future. This thesis attempts to address this gap by answering three main objectives focusing on the quality and utilisation of antenatal care services in Malawi.

The analysis in chapter two examines the impact of the WHO's 2001 FANC model on the utilisation, early access and quality of care in Malawi, using three comparable demographic and health datasets, and the interrupted time series analysis. This method enables the tracking of changes in both levels and the trends of outcome variables. The following three questions are answered in this chapter:

1. What is the impact of FANC on early antenatal care access?
2. What is the impact of FANC on the underutilisation of antenatal care services?
3. What is the quality of the care being offered to pregnant women under the FANC package?

The findings suggest that the adoption of FANC is associated with an improvement in early access to antenatal care. However, the policy has been associated with unintended increases in underutilisation of antenatal care. Furthermore, I found no change in the quality of antenatal care services. Findings from this chapter call for the need to strengthen the health system's capacity through offering more training in FANC guidelines to health providers. Community engagement of health surveillance assistants can also prove effective, especially in rural settings where skilled providers are scarce.

In LMICs, most antenatal care policies are adopted without thorough scientific evaluation and there is lack of empirical evidence on the average number of visits that is likely to produce the most benefit in improving maternal and child health outcomes (Dowswell et al., 2015; Villar et al. 2002). I therefore extend the analysis by estimating the optimal number of antenatal care visits for positive birth outcomes in chapter three. Just like other sub-Saharan African countries, Malawi has struggled to implement the 2001 FANC model effectively. While over 90% of pregnant women access antenatal care at least once with skilled healthcare personnel, only 50% of women received at least four antenatal care visits (NSO, 2016) over a decade since FANC was adopted. The reasons for the implementation failure of FANC have been shown to be due to inadequate equipment, supplies, infrastructure and training (Lungu et al., 2011; Birungi et al., 2006; Nyarko et al., 2006; Chege, 2005). In 2016, the WHO proposed a new model of antenatal care that doubles the minimum number of visits from four to eight based on the

evidence that a reduced model of antenatal care is associated with higher levels of perinatal mortality (Dowswell, 2015). However, if inadequate infrastructure and training are the bottlenecks with the four-visit model, then doubling the minimum number of visits to eight may not improve the situation.

To estimate the optimal number of ANC visits, I used nationally representative Malawi Demographic and Health Survey (DHS) data, and applied instrumental variable models together with highly flexible spline specifications and Wald tests. The application of both instrumental variables models and spline specifications allowed me to estimate breaks in the relationship between the number of antenatal care visits and the probability of low birthweight. In so doing, I establish a threshold at which an additional antenatal care visit has no significant impact in the reduction of low birthweight. I found that only three visits are required to reduce the probability of low birthweight to the same extent as more visits would. The findings from this chapter suggest that low-income health systems are likely to perform just as well if fewer routine visits are conducted with more attention to quality, and reserving additional antenatal care visits to women who critically need them.

The analyses in chapters two and three take advantage of the publicly available nationally representative household Malawi DHSs, which are based on women's self-reports on the services provided. However, the reliability of this data depends on a number of factors including; the client's ability to recall with accuracy; how much attention the client paid to the provider's actions; the client's access to information and knowledge of the content of care; an understanding of the questions being asked and the ability to link them to what the provider was doing; and a willingness to participate, among others (Franco et al., 2002). Furthermore, a client's self-reports may be affected by the current characteristics or circumstances of the individual (anchoring) (Von Fintel & Posel, 2015). The highlighted limitations may lead to an upward or downward bias in the quality of care measurement, which has implications for monitoring the quality of care.

In chapters two and three, an effort was made to reduce recall bias by limiting the samples to the most recent birth prior to the surveys. However, as described, there are other potential biases, which require further exploration. These are addressed in chapter four, where I assessed the extent to which the three most common data-recording methods could contribute to the global and national monitoring of the provision of high-quality maternal healthcare services in

Malawi. Using direct observations during antenatal care consultation as a gold standard, I compared these observations to:

- exit interviews with pregnant women after an antenatal care consultation; and
- retrospective self-reports on the provision of antenatal care services as captured in the demographic health surveys.

I specifically tested the sensitivity, specificity and receiver-operating curves (ROCs) of antenatal care quality indicators as captured in the three survey methods. The results suggest that women tend to overestimate the quality level of service provision. In the client exit interviews, the reliability of women's reports is limited by their lack of knowledge about complicated items of care, for example, counselling on the side effects of iron and the danger signs of complications in pregnancy. This was shown by the different reporting patterns based on the level of education. In the retrospective DHS women's self-reports, I found evidence of recall errors in the women's reports, even among women who had given birth a year prior to the survey. The main recommendation from this chapter is that, in measuring quality of care, it is important to compare women's self-reports with direct observations or facility data in order to get accurate quality estimates.

In the last chapter, chapter 5, I present the key findings and discuss policy recommendations. The overall findings from the thesis suggest that, when it comes to policy implementation and effectiveness, one size does not always fit all. For LMICs, on the supply side, effective maternal health policy interventions require strategies that aim to strengthen the existing structures and reduce system inefficiencies. On the demand side, strategies that increase access to information on the benefits of antenatal care and empower women through education could prove effective.

CHAPTER TWO

Examining the impact of the WHO's Focused Antenatal Care policy on early access, underutilisation and quality of antenatal care services in Malawi: A retrospective study

Abstract

In low- and middle-income countries (LMICs), various antenatal care models have been implemented over the past decades, as proposed by the World Health Organization. One of these models is the 2001 Focused Antenatal Care (FANC) programme. FANC recommends a minimum of four visits for women with uncomplicated pregnancies and emphasises quality of care to improve both maternal and neonatal outcomes. Malawi adopted FANC in 2003; however, up to now no study has been done to analyse the model's performance with regard to antenatal care service quality and utilisation patterns. The methodology in this chapter is based on data pooled from three comparable nationally representative Malawi Demographic and Health Survey (DHS) datasets (2000, 2004 and 2010). The DHS collects data on demographics, socio-economic indicators, antenatal care and the fertility history of reproductive women aged 15–49. I pooled a sample of 8 545 women who had a live birth in the last five years prior to each survey and measured the impact of FANC on early access to care, underutilisation of care and quality of care with interrupted time series analysis. This method allowed tracking changes in both the levels and the trends of the outcome variables. I found that FANC is associated with earlier access to care. However, it has also been associated with unintended increases in underutilisation. I saw no change in the quality of antenatal care services. In light of the WHO 2016 antenatal care guidelines, which recommend an increase of visits to eight, these results are important. Given that I find underutilisation when the benchmark is set at four visits, eight visits are unlikely to be feasible in low-resourced settings.

2.1   Background

In low and middle-income countries (LMICs), various antenatal care models have been implemented over the past decades to improve both maternal and child health outcomes, as proposed by the World Health Organization (WHO) (Villar et al., 2001). One of these models is the 2001 Focused Antenatal Care (FANC) programme. In this chapter, I consider whether FANC contributed towards maternal health by improving early access, increasing the number of visits and enhancing the quality of care. The WHO began promoting FANC in 2001, replacing the traditional antenatal care service model, which included numerous antenatal visits (7–16 visits) and had proved to be a challenge in resource-constrained settings (WHO, 2002a).

FANC recommends only four antenatal care visits for women with uncomplicated pregnancies, and more otherwise. The four-visit model emphasises quality of care and provides a package of services that contributes to the health and well-being of women during pregnancy, childbirth and the post-delivery period (WHO, 2002a). The four visits in the FANC model are scheduled to be made at specific times, as follows: the first visit should occur between 8 and 12 weeks after conception but not later than 16 weeks; and a further three visits should occur between 24 and 38 weeks of gestation (WHO, 2002a).

Most LIMCs, including Malawi, incorporated FANC into their healthcare systems (Villar et al., 2001). Despite the WHO revision of antenatal care guidelines to double the number of visits from four to eight in 2016, FANC is the antenatal care model currently in use in Malawi (Mamba et al., 2017). Evidence shows that the model had substantial public health implications, especially in low-income countries where healthcare resources are inadequate (Villar et al., 2001). Moreover, in the 2015 Cochrane review (Dowswell et al., 2015), it was argued that the reduced-visit model reduces the costs for women. This includes commuting times to and from clinics, waiting time, transport costs to clinics located far away, loss of hours from work, and care of other children at home (Dowswell et al., 2015). However, little is known about the impact of FANC on the early access to, and the utilisation and quality of antenatal care services in sub-Saharan African countries – including Malawi (MoH, 2007). This study seeks to fill this research gap.

Previous research on the impact of FANC has been inconclusive. Trials conducted by the WHO in Argentina, Cuba, Saudi Arabia and Thailand in 2001 showed that FANC was safe and could be easily maintained, comprehensive and cost-effective antenatal care model (WHO, 2002a).

14

In Kenya, the adoption of FANC led to improved detection of existing diseases in pregnancy during the first antenatal care visit, planning for birth, prevention of complications, and postpartum counselling (Birungi & Onyango-Ouma, 2006). In Ghana, FANC resulted in improved quality and continuity of care (Nyarko et al., 2006). There are exceptions, however. In South Africa, FANC had no significant effect on the quality of antenatal care services. This was attributed to a lack of training, high staff turnover and inadequate supervision (Chege, 2005). In the 2015 Cochrane review study, results show low satisfaction levels with the reduced-visit model by women in both low- and high-income settings and perceived the gap between the scheduled visits as too long (Dowswell et al., 2015).

In Malawi, there is tentative evidence, based on research at one site, which shows that the introduction of FANC led to improvements in the quality of antenatal care services at the facility (Lungu et al., 2011). Since the study used only one urban clinic, its external validity is questionable. This study expands on that work by considering a nationally representative sample of women who accessed antenatal care in clinics across Malawi, including rural clinics, to assess the effectiveness of FANC at a national level.

The study also adds to the literature on the impact of FANC in African countries by incorporating a time dimension in the analysis. I pooled three cross-sectional DHS datasets and use the year of the mother's delivery as the date stamp (instead of the survey year, see more details in Table 2.1). This is unlike previous studies that used one cross-sectional study at a point in time to look at correlations and therefore are not suited to provide statistical evidence on the effectiveness of FANC policy. Furthermore, the interrupted time series methodology allows the tracking of changes in both the levels and the trends of the outcome variables.

Malawi is also an important case study because of its challenging policy implementation environment arising from high levels of maternal mortality (GoM, 2014), high levels of poverty (World Bank, 2015), a lack of skilled medical personnel (WHO, 2017) and a lack of health infrastructure (World Bank, 2015; Lungu et al., 2011). In resource-constrained settings like this, it is important to understand whether a reduction in the number of visits and stricter guidelines about the content of each visit could improve the quality of care and thus enhance maternal and child health. This study engages with the debate on the appropriate model of care and the recently proposed WHO reforms to increase the recommended antenatal care visits to eight.

2.1.1 Study context

Malawi is classified as a low-income country with a GDP per capita as low as $274 in 2014. This translates into $0.75 that the average individual can spend per day (World Bank, 2015). Given the low GDP, the government has a limited tax base and faces dramatic trade-offs in its policy decisions while having to deal with considerable need (GIZ Health, 2011). These challenges are further exacerbated by a healthcare system with poor infrastructure, a lack of equipment and qualified human resources, and weak management (WHO, 2017).

In Malawi's healthcare system, services are delivered at primary, secondary and tertiary levels (MoH, 2014b) which are linked through an elaborate referral system (WHO, 2008a). Primary healthcare is the lowest tier of care and consists of maternity units, health centres, village clinics, health posts, dispensaries, and community and rural hospitals (MoH, 2014b). The second tier, the secondary level, constitutes district hospitals and provides specialised services to patients referred from the primary healthcare level through either outpatient or inpatient services (MoH, 2014b). The specialised services are enhanced by support services, such as diagnostic, laboratory, rehabilitation, blood bank and physiotherapy services. The third tier, tertiary healthcare services, consists of highly specialised services and constitutes central and other specialist hospitals.

Only 65% of facilities in Malawi offer antenatal care services, including government, non-profit and private providers (MoH, 2014b). In government facilities, the provision of antenatal care services is integrated with under-five clinics, postnatal care, family planning, and other reproductive health services and is provided free of charge (MoH, 2014b). For-profit and non-profit providers (such as the Christian Health Association of Malawi (CHAM)) require user fees at the point of use. As of 2010, 73% of antenatal care services were provided at primary health facility level on a daily basis while 27% were provided at secondary and tertiary levels (MoH, 2014b). In Malawi, nurses and midwives provide 80% of antenatal care services, whereas the rest of the population receive antenatal care from clinical officers, doctors or maternal-child health aides (MoH, 2014b).

According to the Ministry of Health report (2007), Malawi adopted the FANC policy in 2003, and was therefore, integrated in the essential health care package (EHP). JHPIEGO and WHO/Malawi provided financial as well as technical support to the Ministry of health for an in- service training of service providers in the new FANC guidelines with the goal of

stimulating change (MoH, 2007). At community level, sensitization activities to communicate changes in the delivery of antenatal services were undertaken to pregnant women through health talks and individual messages given by healthcare workers (MoH, 2007). Given these changes, I expect that the introduction of FANC improved the quality of ANC services as well as increased the utilisation of ANC services. I test this hypothesis in this chapter.

The successful implementation of FANC required that healthcare facilities have adequate clinical skills, infrastructure, essential equipment, medication and laboratory supplies (Lungu et al., 2011). However, the government of Malawi did not take the necessary measures to invest in the resources required for the successful implementation of FANC. In the 2010 annual monitoring report on the implementation of FANC showed that only one of the four central hospitals and four of the 24 district hospitals in Malawi met the WHO standards for delivering FANC (MoH, 2010). Moreover, as of 2014, the WHO reported that, there were only 0.2 doctors and 3.4 nurses and midwives for every 10 000 people in Malawi (WHO, 2017).

2.2 Methods

2.2.1 Data

This study is a retrospective study. It uses three Malawi Demographic and Health Survey (DHS) datasets conducted in 2000, 2004/2005 and 2010. I chose the three years based on comparability. The DHS provides comprehensive health information for women of reproductive ages between 15 and 49 and their children. The survey uses a multi-stage cluster sampling design to select households for participation based on the Malawi population censuses of 1998 and 2008. Random sampling of the enumeration areas and household listing operation is conducted in the first stage, which is then followed by a random sampling of the households in the second stage. The study uses the women data file of the DHS, which contains data on, among others: demographics, household socio-economic status, and antenatal care utilisation practices. Information on the utilisation of antenatal care services and the components of care is reported on women who had a live birth during the five years before each survey. The response rate for each survey was above 95%.

The main independent variable of interest in this study is the FANC policy dummy, which captures the year when FANC was implemented in Malawi (see Table 2.1 for details). I created the FANC policy dummy using the mother's year of delivery or the child's year of birth. All women who gave birth after 2003 were categorised to be in the post-FANC period and those

who delivered prior to 2003 were in the pre-FANC period. The total initial pooled sample was 28 763 women; however, I limited the analysis to women who delivered three years before and three years after the adoption of FANC, excluding women who gave birth in 2003. This restriction reduced the final sample to 8 545 women. Limiting the analysis to the years that were closest to the launch of the policy prevents the influence of other policies introduced prior to 2000 or after 2006, including the 2007 ban on traditional birth attendants (TBAs) (MoH, 2007).

Table 2.1: Cross-tabulation of birth year and FANC

DHS Data enumeration year	Woman's year of delivery	FANC	Frequency
2000	1995	0	102
2000	1996	0	512
2000	1997	0	1 023
2000	1998	0	1 860
2000	1999	0	2 494
2000	2000	0	2 386
2004/2005	2001	0	791
2004/2005	2002	0	1 352
2004/2005	2003	1	2 308
2004/2005	2004	1	2 240
2004/2005	2005	1	428
2010	2006	1	1 348
2010	2007	1	2 247
2010	2008	1	3 545
2010	2009	1	3 924
2010	2010	1	2 203
Total sample size			**28 763**

Source: Malawi DHS 2000, 2004 and 2010

Notes: 0 = FANC policy not adopted; 1 = FANC policy adopted

2.2.2   Main outcome measures

The research considers the impact of FANC on three outcomes: early access to antenatal care, inadequate use/underutilisation of care, and quality of care.

Early access to care

In this study, I use the timing of first antenatal care visit to measure early access. This variable ranges from 0 to 9 months. A Weibull hazard model (an example of a survival-analysis model) was used to model the gestational age at which the mother enters the antenatal care system. Weibull models were initially developed to consider the survival of machine components, so that interpretation is sometimes counterintuitive: a higher likelihood of earlier 'component

18

failure' here is equivalent to a higher likelihood of an antenatal care visit at an earlier gestational age. Survival-analysis models are common in the health sciences (Zhu et al., 2011; Naomi et al., 2009; Alan, 1980). I study differences in early access to care before and after FANC was implemented.

Underutilisation of care

FANC recommends a minimum of four antenatal care visits for women with uncomplicated pregnancies, therefore, in this context, a woman with fewer than the minimum number of four visits has underutilised the services. For the model, I sought a more precise definition of underutilisation that does not overlap with that of early access. I therefore limited the sample to women who initiated their first antenatal care visit by 16 weeks of pregnancy as required by FANC to avoid duplication and overlap with the early-access indicator. By limiting the analysis to women who accessed antenatal care early, I avoided a 'double count' problem where underutilisation may merely be another manifestation of late access. The definition also excludes women who never initiated ANC.

Underutilisation is a binary variable defined as 1 if a pregnant woman initiated her first visit in the first trimester of pregnancy but did not make the recommended number of four visits, and 0 otherwise. For the model, I considered cluster averages of underutilisation for the subset of women who initiated care early. Underutilisation is thus the likelihood of women in a specific cluster underutilising antenatal care service provided that they had a first visit by 16 weeks of pregnancy as prescribed by FANC. I analyse this outcome with ordinary least squares (OLS) regressions.

Quality of care

The aim of FANC is to achieve not only a minimum number of four visits but also compliance with FANC protocols. I therefore also track whether healthcare workers complied in conducting eight key antenatal care tests or examinations. These include routinely conducted diagnostics (taking blood and urine samples), physical examination (measuring blood pressure and weight), and other preventive procedures (administration of tetanus toxoid, prophylaxis, iron and folic supplements and establishing complication readiness). These questions were asked in each of the DHSs and are here interpreted as proximate indicators of the quality of antenatal care services.

As was the case for underutilisation, the analysis sample was also limited to women who initiated antenatal care early. The first visit recommended by FANC is critical for HIV-infected pregnant women, to ensure access to antiretroviral (ARV) prophylaxis. In addition, pregnant women can access early interventions such as syphilis screening and treatment, provision of iron supplements and malaria prevention and treatment (MoH, 2003). When a woman accesses her first antenatal care visit later than 16 weeks of pregnancy, she risks missing some of the early antenatal care interventions crucial for early detection of pregnancy complications. In such a case, missing some required interventions would therefore not represent non-compliance with FANC protocols, but result from late access.

Using the identified eight key measures, an antenatal care composite-compliance index was constructed using multiple correspondence analysis (MCA) (Chama-Chiliba, 2015; Greenacre & Blasius, 2006). Table 2.2 presents the variables included in the composite index, with their respective categories and weights. The weights are identified from the first dimension of the MCA, which explained about 85% of the total inertia. Positive (negative) weights reflect the higher (lower) quality of care.

Table 2.2: Results for multiple correspondence analysis for the ANC quality index

Variable	Categories	Weights
Prophylaxis	Given malaria prophylaxis	0.417
	Not given malaria prophylaxis	-2.103
Blood pressure	Blood pressure measured	0.597
	Blood pressure not measured	-2.662
Blood sample	Blood sample taken and tested for disease	1.080
	Blood sample not taken or tested	-1.576
Urine sample tested	Urine sample taken and tested	1.927
	Urine sample not taken or tested	-0.557
Iron tablets	Iron tablets/syrup given	0.325
	Not given iron tablets/syrup	-1.789
Weight measured	Weight measured during visit	0.174
	Weight not measured	-5.196
Complications	Told about complications	0.514
	Not told about complications	-1.612
Tetanus toxoid	Given at least one tetanus toxoid vaccine	0.265
	Not given any tetanus toxoid vaccine	-1.686

The ANC-compliance index was then used as the main outcome variable in the ordinary least squares regression to estimate the statistical impact of FANC on the quality of antenatal care services. Again, I aggregated at the community level. This shifted the attention from the

demand side (women's individual experiences) to the supply side (what can be expected of a community clinic to provide on aggregate).

2.2.3 Distinguishing FANC from confounding influences over this period

The major concern in this study is distinguishing the effects of FANC from two other shifts that occurred during this period: the expansion of education; and the rise in HIV prevalence and the launch of the prevention of mother-to-child transmission of HIV (PMTCT).

Women's education level

Women's education level has been included categorically with values ranging from 0 to 3 (0, no education; 1, primary education; 2, secondary education; and 3, tertiary education). Free primary education (FPE) in Malawi was implemented in 1994 (Al-Samarrai & Zaman, 2018). I expect post-FANC mothers to be more educated than pre-FANC mothers. Higher educational attainment is associated with a higher likelihood to use maternal healthcare services (Sagna & Sunil, 2012; Gage, 2007; Grossman, 1972), thus, it is important to include educational attainment in the models.

HIV testing and the introduction of PMTCT

In order to reduce the prevalence of HIV/AIDS, Malawi launched the national prevention of mother-to-child transmission (PMTCT) programme in 2001 to prevent the transmission of HIV from an HIV positive mother to her infant during pregnancy, labour, delivery, or breastfeeding(WHO, 2014b). The main concern was that the promotion of PMTCT may have caused women to access care earlier and could have affected the quality of these services. To distinguish the impact of FANC from that of the PMTCT programme, I included the community-level proportion of pregnant women who reported ever taking an HIV test.

2.2.4 Other control variables

Pregnancy risk factors

As per the FANC guidelines (WHO, 2002a), women with uncomplicated pregnancies are recommended to make four antenatal care visits. However, women with a history of any complications or illnesses in previous pregnancies are identified to be high risk and recommended to make more than four visits. In this study, I included the following variables as proxies for a risky pregnancy: age at birth of less than 16 and more than 40, a history of

21

miscarriage and/or Caesarean deliveries. These indicators were chosen based on FANC guidelines (WHO, 2002a) and data availability.

Type of health worker

In the analysis of the impact of FANC on the quality of care, I also controlled for the type of health worker providing the antenatal care service to the client. I expect skilled healthcare providers such as doctors and nurses to be more knowledgeable and equipped to provide quality service than a ward attendant. Studies in Nigeria (Fagbamigbe & Idemudia, 2015) and Nepal (Joshi et al., 2014) reported that skills acquired by antenatal care providers have positive significant impacts on the quality of antenatal care offered. I defined this variable as a series of binary variables (doctor; nurse; traditional birth attendant and health surveillance assistant/ward attendant), as captured in the DHS dataset. However, I excluded one category, 'Other', to avoid the dummy variable trap.

Other factors

I expect rural-based women to have poor access to clinics due to the longer travel distance to the closest facility. I also expect women who are exposed to media (such as radio or television) to have more health-related information and therefore to utilise antenatal care services more regularly. Other studies have also found that media penetration through radio is important in influencing the use of maternal healthcare (Titaley et al., 2010; Simkhanda et al., 2008). I could unfortunately only capture these traditional forms of media access, as the DHS did not include questions on social media or internet access. Even more so, supplementary statistics[2] indicated that these channels were not an important means through which FANC could have been promoted during the implementation period. It should therefore not be an important channel of influence for FANC services uptake over the period of my analysis. By contrast, close to half of the women in the sample reported that they listen to the radio every day and that radio was part of the government's promotion campaign for FANC (MoH, 2007).

2.2.5   Estimating the impact of FANC

I used interrupted time series analysis (ITSA) applied to ordinary least squares or Weibull models to estimate the changes in the outcome variables of interest between the two periods. ITSA captures changes both in levels and in trends. The ITSA design is mainly used to evaluate

---

[2] The second Malawi household integrated survey estimated that, between 2004 and 2005, only 3.16% of the households in Malawi and 3.52% of women of reproductive age had access to mobile phones (NSO, 2005).

interventions that are implemented at a population level over a distinct time period (Soumerai, 2015; Bonell et al., 2009; Victoria et al., 2004). In standard ITSA, the segmented-regression model is defined as follows:

$$Y_{it} = \beta_0 + \beta_1 T_{it} + \beta_2 X_{it} + \beta_3 T_{it} X_{it} + \beta_4 Z_{it} + \mu_{it} \tag{1}$$

Where:

Y_{it} is one of the outcomes of interest (in this case, quality of ANC services index; timing of first ANC visit; proportion of women with inadequate visits/underutilisation);

i indexes either the individual or geographic cluster;

t indexes time or year of delivery;

β_0 represents the baseline level of the outcome at $t = 0$;

β_1 is interpreted as the change in outcome associated with a time-unit increase (representing the underlying pre-intervention trend);

β_2 captures any discrete changes following the intervention;

X is the FANC dummy variable indicating the pre-intervention period (coded 0) or the post-intervention period (coded 1);

T measures the time before (negative) or after (positive) the intervention, and is included as a linear trend;

β_3 indicates the slope change following the intervention (using the interaction between time and intervention: TX);

Z is a vector of controls for confounding factors; and

μ is the error term.

The study's main interest lies in assessing whether the time progression has been altered by FANC, so that is the coefficient of interest. All statistical analyses were conducted in Stata Version 14.

2.3 Results

2.3.1 Descriptive analysis

Descriptive statistics in Table 2.3 illustrate changes between the pre- and post-FANC periods in some of the variables used in the analysis. The results show significant increases in the proportion of women who accessed ANC by 16 weeks of pregnancy and quality of antenatal care services in the post-FANC period. However, in the post-FANC period, the average number

23

of antenatal care visits declined to just below the minimum number of four visits required by FANC.

The proportion of women with no education is lower in the post-FANC period at 21% compared to the pre-FANC period at 27%. The proportion of women who reported having been tested for HIV between the two periods jumped from 15% before FANC to 41% after FANC. I saw a statistically significant increase in two of the pregnancy risk factors: the proportion of women with a history of Caesarean delivery increased from 3% to 5% after FANC, and miscarriage increased from 11% to 13% after FANC. Lastly, there was also a significant increase in media access through radio and television in the years following FANC implementation.

Table 2.3: Social and demographic characteristics

Variables	Means or proportions		P-value
	Pre-FANC	Post-FANC	Difference in means or proportions
Main outcome variables			
Early access	0.26	0.29	0.0007***
Number of visits	4.04	3.76	0.0000***
Standardised ANC quality index	0.71	0.74	0.0000***
Confounding variables			
Mother's education level			
No education	0.27	0.21	0.0000***
Primary	0.62	0.65	0.0369**
Secondary	0.10	0.14	0.0000***
Tertiary	0.00	0.01	0.0291**
Type of residence			
Rural	0.85	0.84	0.0013**
Urban	0.15	0.14	0.0013**
HIV-test statistics			
Women who reported having ever been tested for HIV/AIDS	0.15	0.41	0.0000***
Pregnancy risk indicators			
Age: 16 and below	0.02	0.02	0.6901
Age: above 40	0.04	0.05	0.1182
History of miscarriages	0.11	0.13	0.0059*
History of Caesarean deliveries	0.03	0.05	0.0004***
Media access			
Radio			
Not at all	0.23	0.22	0.3416
Less than once a week	0.18	0.16	0.0014**
At least once a week	0.12	0.16	0.0000***
Almost every day	0.47	0.46	0.2414

	Means or proportions		P-values
	Pre-FANC	**Post-FANC**	**Difference in means or proportions**
Television			
Not at all	0.90	0.83	0.0000***
Less than once a week	0.05	0.08	0.0000***
At least once a week	0.02	0.03	0.0000***
Almost every day	0.03	0.06	0.0000***

*Notes: ***1% level of significance; ** 5% level of significance; and *10% level of significance.*

2.3.2 Early access to care

Figure 2.1 shows changes in the timing of the first antenatal care visit between pre- and post-FANC periods. The *y*-axis represents the month after conception during which a woman initiated her first antenatal care visit and the *x*-axis represents the number of years from the implementation of FANC. Negative (positive) values are the number of years before (after) FANC implementation, while zero indicates the time of implementation in 2003.

Figure 2.1: Timing of first ANC visit
Notes: lpoly smoother, no controls – X-axis; 0 – represents 2003, the year of FANC implementation; vertical line demarcates the two periods. Negative years indicate number of years before FANC and positive years the number of years after FANC implementation. Y-axis – timing of first ANC visit (measured in months).

Similar to the descriptive statistics, I noticed a significant reduction in late access following FANC implementation, over and above the existing downward trend that emerged before FANC. Table 2.4 presents the Weibull regression results for the relationship between FANC adoption and early access to care. I found that FANC had decreased the likelihood of early access when I look merely at the levels, but this effect is overwhelmed by the positive trend effect. It is encouraging to see that there is a positive coefficient on the post-FANC trend and that the implementation of this policy has increased early access. Furthermore, having at least primary-level education, access to radio, a history of miscarriage or caesarean delivery, urban residence and having been tested for HIV are all associated with a higher likelihood of early access to care.

Table 2.4: Impact of FANC on early access to care: Weibull model with interrupted time series analysis

Variables	Coefficient	Standard error
FANC policy dummy		
FANC	-0.370***	(0.086)
Time trend	0.018	(0.031)
Post-FANC x Time trend	0.207***	(0.036)
Woman's education level (base=no educ)		
Primary	0.192***	(0.042)
Secondary	0.363***	(0.066)
Tertiary	0.463**	(0.201)
Pregnancy risky indicators		
16 years and below	-0.156	(0.141)
Above 40	-0.218***	(0.074)
Miscarriage	0.108***	(0.042)
Caesarean deliveries	0.126	(0.077)
Access to radio (Base=no access)		
Less than once a week	0.110**	(0.051)
At least once a week	0.078	(0.055)
Almost every day	0.204***	(0.044)
Access to television		
Less than once a week	-0.186	(0.061)
At least once a week	0.047	(0.109)
Almost every day	0.051	(0.077)
Community type		
Urban	0.110**	(0.046)
Community HIV-test statistics	0.105	(0.089)
Constant	-7.182***	(0.0125)
Observations	**8 522**	
Wald Chi2(1)	**299.92*****	

*Notes: Weibull survival models. ***1% level of significance; ** 5% level of significance; and *10% level of significance. Coefficients with standard errors in parentheses.*

2.3.3 Underutilisation of care

Figure 2.2 shows an increase in the community-level underutilisation of antenatal care services after FANC adoption. Table 2.5 confirms that FANC has had the unintentional consequence of increasing underutilisation of antenatal care. Before FANC, women were progressively less likely to underutilise antenatal care services over time. After FANC was adopted, underutilisation increased by eight percentage points per year (after accounting for the confounding variables), relative to the initial downward trend. Very few of the control variables and the proxies for confounding factors were statistically significant in these regressions. Those respondents who had access to radio at least once a week were less likely to underutilise care.
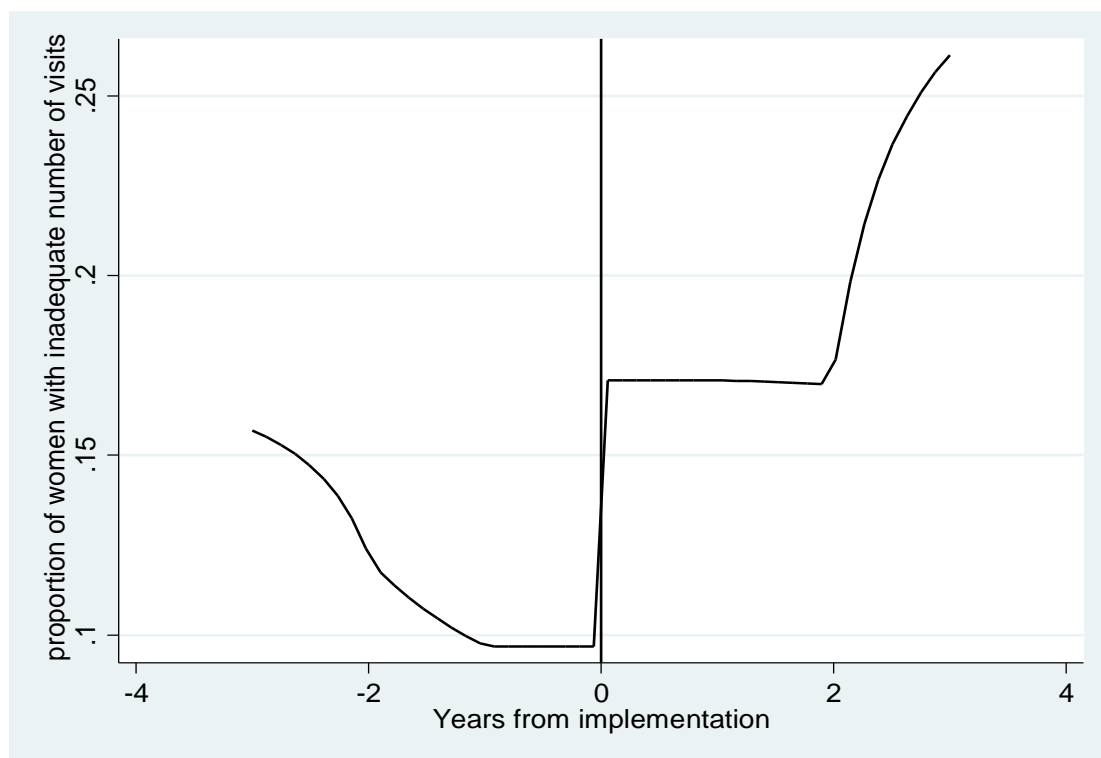
Figure 2.2: Change in underutilisation of ANC services in a cluster
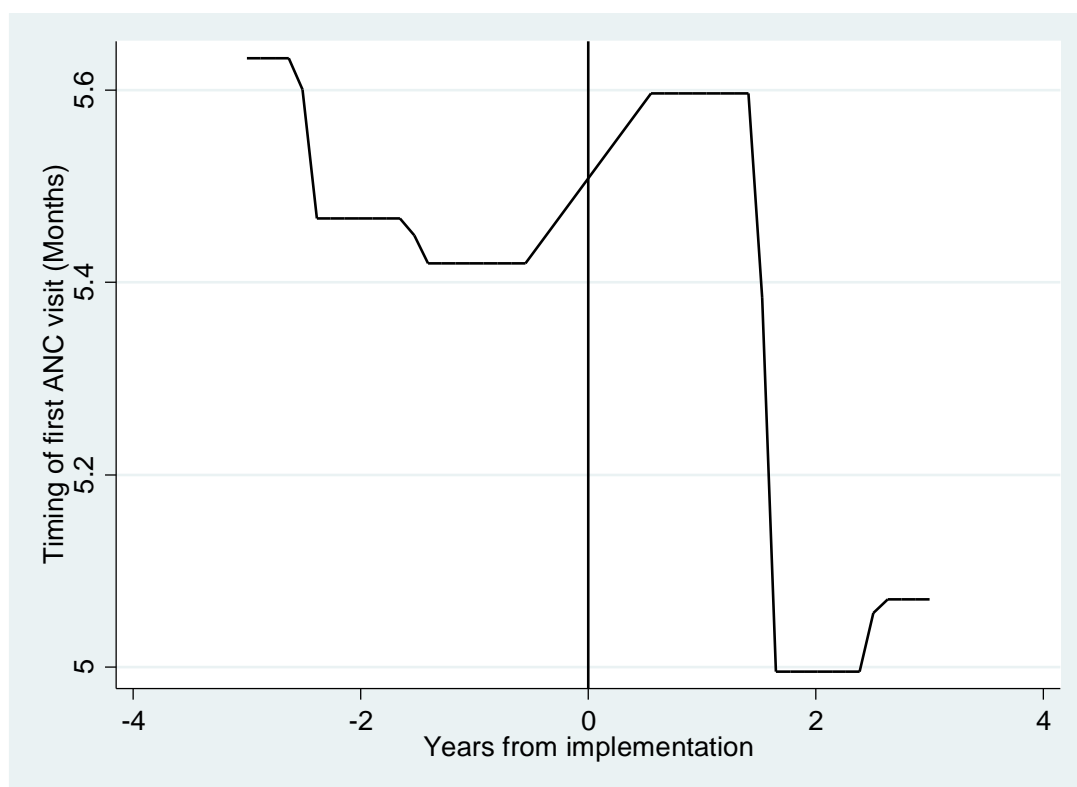
Notes: lpoly smoother, no controls – X-axis; 0 – represents 2003, the year of FANC implementation; vertical line demarcates the two periods. Negative years indicate number of years before FANC and positive years the number of years after FANC implementation. Y-axis – proportion of women with inadequate number of visits in a cluster.

Table 2.5: Impact of FANC on underutilisation of ANC services: OLS model with interrupted time series analysis

Variables	Coefficient	Standard error
FANC Policy dummy		
FANC	0.034	(0.046)
Time trend	-0.028**	(0.014)
Post-FANC x Time Trend	0.083***	(0.021)
Woman's education level (Base= No educ)		
Primary	0.016	(0.031)
Secondary	0.068	(0.044)
Tertiary	0.106	(0.144)
Media access (Base=No access)		
Radio		
Less than once a week	-0.081*	(0.042)
At least once a week	-0.103**	(0.044)
Almost every day	-0.038	(0.034)
Location		
Urban	-0.025	(0.027)
HIV test	-0.038	(0.031)
Television	Yes	
Pregnancy risky indicators	Yes	
Constant	0.119**	(0.046)
Observations	1 606	
R-squared	0.0406	

Notes: Ordinary least squares regression. ***1% level of significance; ** 5% level of significance; and *10% level of significance. Coefficients with standard errors are in parentheses.

2.3.4   Quality of care

Table 2.6 reports descriptive results on individual-level quality of care. The findings show an overall improvement in compliance of individual antenatal care contents after FANC. Women who reported to have received all eight components increased from 9% to 11% in the post-FANC period. There were significant increases in the taking of blood and samples, communication about complications, dispensing of iron tables and dispensing of malaria prophylaxis.

In Figure 2.3, I show the change in overall quality of antenatal care services at the community level, as measured by the MCA index. In the years immediately after FANC implementation, there was a decline in quality, which was then followed by a period of stagnation before a substantial improvement. This could represent a one-year lag or initial teething problems with the implementation of this policy.

Table 2.6: Services received during ANC visit

Variable categories	Proportion of women		P-value
	Pre-FANC	Post-FANC	
Weighed during pregnancy	0.96	0.95	0.2396
Blood pressure checked during pregnancy	0.79	0.79	0.4546
Urine sample taken during pregnancy	0.20	0.22	0.0638*
Blood sample taken during pregnancy	0.37	0.47	0.0000***
Told about complications	0.69	0.72	0.0032***
Given or instructed to go buy iron tablets	0.77	0.84	0.0000***
Took Fansidar as prophylaxis for malaria	0.76	0.84	0.0000***
Received at least one tetanus injection	0.83	0.83	0.3319
Received all eight components	**0.09**	**0.11**	**0.0002***

*Notes: ***significant at 1% level of significance; ** 5% level of significance; and *10% level of significance.*
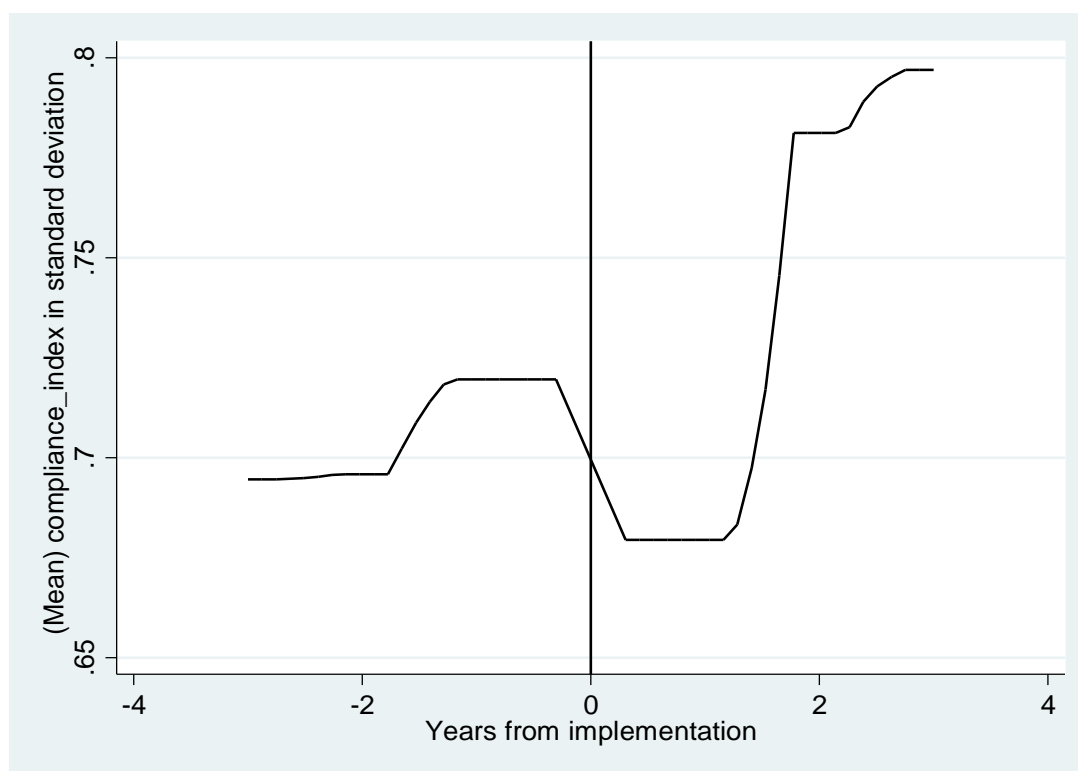
Figure 2.3: Change in quality of care

Notes: lpoly smoother, no controls – X-axis; 0 – represents 2003, the year of FANC implementation; vertical line demarcates the two periods. Negative years indicate number of years before FANC and positive years the number of years after FANC implementation. Y-axis – cluster level mean compliance index in standard deviation.

The coefficients reported in Table 2.7 show no evidence of FANC improving the quality of care. Based on the descriptive results, I consider the hypothesis that there may have been

teething problems with implementation or a lag in implementation, changing the policy cut-off point to 2004, but this specification also does not provide any evidence of an improvement in the quality of care. Urban location, education and HIV-testing prevalence rates are all positively associated with quality of care. In addition, skilled health providers (doctors and nurses) are more likely than unskilled providers (traditional birth attendants) to comply with the set of protocol measures in the quality index.

Table 2.7: Impact of FANC on the quality of ANC services: OLS model with interrupted time series analysis

Variables	Coefficient	Standard error
FANC policy dummy		
FANC	-0.067***	(0.020)
Time trend	0.013**	(0.006)
Post-FANC x Time trend	0.008	(0.009)
Woman's education level (Base=No educ)		
Primary	0.042***	(0.013)
Secondary	0.096***	(0.018)
Tertiary	0.023	(0.062)
Community type		
Urban	0.032***	(0.011)
HIV-testing prevalence		
HIV-test statistics	0.086***	(0.013)
Pregnancy risk indicators		
16 years and below	-0.023	(0.043)
Above 40 years	0.031	(0.024)
Miscarriage	0.001	(0.015)
Caesarean	0.018	(0.023)
Type of health worker		
Doctor	0.043**	(0.020)
Nurse	0.062***	(0.021)
Traditional birth attendant	-0.324***	(0.033)
Ward attendant / health surveillance assistants	-0.002	(0.023)
Constant	0.632***	(0.027)
Observations	**1 606**	
R-Squared	**0.2252**	

*Notes: Ordinary least squares regressions. ***significant at 1% level of significance; ** 5% level of significance; and *10% level of significance. Coefficients with standard errors in parentheses.*

2.4   Discussion

Antenatal care provides a valuable opportunity to support pregnant women and ensure that both them, and their babies, profit from effective and good-quality maternal care. This chapter analysed the impact of the 2001 FANC model on the quality and use of antenatal care services currently provided in Malawi. Unlike previous studies, this study used nationally representative

data and appropriate statistical methods to provide evidence of the impact of FANC in low-resourced settings. Given the 2016 revision of antenatal care guidelines, this study is timely and important in order to evaluate the effectiveness of the FANC policy, which is still being followed in Malawi.

2.4.1 Early access to care

The results show a positive association between FANC and early access to antenatal care. The findings demonstrate a significant increase in the proportion of women who accessed ANC by 16 weeks of pregnancy of about 21 percentage points per year (relative to baseline) in early access after FANC adoption. FANC has helped to promote early access to antenatal care in Malawi. I find the results plausible because FANC required the first contact to occur by 16 weeks of pregnancy (WHO, 2002a). However, the descriptive results show that three years after the implementation of FANC, only 29% of women in the sample had initiated antenatal care by 16 weeks of pregnancy. At a national level, the DHS reports that only 24% of women initiated antenatal care early in 2015 (NSO, 2016). Many women delay access to care beyond the first trimester. This is worrying, given that HIV transmission remains a concern.

Furthermore, women who have at least primary education are more likely to initiate their first antenatal care visit by 12 weeks of pregnancy. This finding is similar to that of Yaya et al. (2017), who argued that educated women are more knowledgeable about the significance of antenatal care visits. Similarly, a study conducted by Furuta and Salway (2006) in Nepal suggests that education promotes novel values and attitudes that are favourable to the use of modern healthcare, thereby increasing the likelihood of women accessing skilled care. Consequently, empowering females through education may be effective in increasing uptake of antenatal care services are likely to also increase the use of antenatal care (Joshi et al., 2014). This is particularly important in Malawi, where only one in four women have at least some secondary education and only 3% of women have tertiary education (NSO, 2016).

Women with a history of miscarriage and/or Caesarean deliveries are also more likely to initiate antenatal care early, confirming the results found by Pell et al. (2013). This finding is of particular importance because it shows that women are responding to FANC's emphasis on complications and danger signs to watch for in pregnancy. The study also finds that women who have access to a radio at least once a week are more likely to initiate their first antenatal care visit by 12 weeks of pregnancy. This finding is consistent with findings by Bbaale (2011)

who found a positive relationship between increased media campaigns and utilisation of antenatal care services in Uganda.

The type of residence a woman reports to come from has a significant impact on early antenatal access in this study. Women who reside in urban communities were more likely to initiate their first antenatal care visit early compared to women from rural communities. This finding is consistent with previous studies (Kawungezi et al., 2015; Gudayu et al., 2014; Gupta et al., 2014). Supply factors such as the quality of antenatal care services provided and the distance to the healthcare facility have been reported to play an important role (Roberts et al. 2016). Malawi has promoted FANC via free public provision; however, the majority of healthcare facility especially higher-level facilities are located in urban areas urban areas (MoH, 2007).

Other factors influencing the timing of first antenatal care visits, as reported by qualitative studies in Malawi, include women's beliefs (Roberts et al., 2016) and user fees in CHAM hospitals (Mamba et al., 2017). Evidence shows that mothers often seek care or pregnancy lessons from traditional health workers and elderly women before seeing medical personnel at a facility in Mangochi, Malawi (Roberts et al., 2016). This practice stems in part from the belief that it is not necessary to attend clinics or hospitals unless one is physically ill. Furthermore, while FANC health services are free at the point of use in public facilities, in communities where the majority of facilities are managed by CHAM, people pay user fees at the point of use of MKW2 000 (US$3) on average to access FANC services (Mamba et al., 2017). Many rural and poor women in Malawi cannot afford the cost to access private FANC services.

2.4.2 Underutilisation of care

The results show that FANC has been associated with unintended increases in underutilisation of antenatal care services. These results are consistent with a previous Zambian study where the researcher reported a significant decline in the number of visits to an average of three following FANC adoption (Chama-Chiliba, 2014). A qualitative study in Malawi reported that some FANC requirements served as barriers to accessing and adequately utilising FANC services in Mangochi, Malawi (Mamba et al., 2017). In particular, FANC encourages men to engage in reproductive health and preparation for birth. The study reports that most of the healthcare facilities in the district have laws formulated by health officials and community leaders stating that a pregnant woman can only be attended to if her spouse accompanies her. Furthermore, due to the requirement for birth preparedness as suggested by FANC guidelines,

it is mandatory for a woman to bring with them cloth wraps for the baby (traditional 'chitenje') at their initial FANC visit as commanded by the healthcare providers (Mamba et al., 2017). All these requirements reportedly make FANC services expensive and therefore less accessible.

Another interesting result is the negative significant coefficient between underutilisation and access to radio. These results are consistent with those found by Bbaale (2011). Media access, especially access to radio, has been shown to be instrumental in changing behaviour regarding maternal health services utilisation (Yaya et al., 2017; Gupta et al., 2014).

2.4.3   Quality of care

There was an increase in blood testing, but three years after the implementation of FANC, only 47% of women had received blood tests. This level of access is disappointingly low given that blood sample testing is one of the essential factors in reducing HIV transmission from the mother to the unborn child. Under reporting is a distinct possibility due to potential recall bias or the sensitive nature of this question.

Only 22% of women who gave birth after FANC was implemented reported having been given a urine test during the antenatal care visit. Lungu et al. (2011) found that there were shortages of essential laboratory supplies at one urban clinic in Malawi, which made it difficult to conduct the essential laboratory tests. Nyarko et al. (2006) reported similar findings in Ghana. The results also show that only 11% of users in the sample received all the antenatal care components after the implementation of FANC. In Nigeria, only 5% of users received the optimal level of care (all 10 components) after FANC adoption (Fagbamigbe & Idemudia, 2015).

The regression results show no significant improvement in the quality of antenatal care services, even after accounting for a one-year lag in adequate implementation of the policy. This is similar to the experience in South Africa, where FANC did not improve performance in maternal health delivery (Chege, 2005). A multi-country-level study conducted in selected LMICs found very low compliance rates with FANC standards among women who went for their required four antenatal care visits (Benova et al., 2018). The results of these studies show that LMICs struggled to implement FANC and, as such, the policy did not translate into improved quality in antenatal care services.

The results did, however, show a significant positive association between HIV-testing share and quality index, suggesting that PMTCT may have had an impact on quality. I also found that the type of health worker providing antenatal care has an important role in determining the quality of antenatal care received by women. Women were more likely to receive good-quality antenatal care if skilled providers (doctors or nurses) attended to them compared to those who were relatively less skilled (patient attendant / health surveillance assistants and traditional birth attendants). Similar results were also reported in Nigeria (Fagbamigbe & Idemudia, 2015) and Nepal (Joshi et al., 2014). However, a qualitative study analysing women's perspectives on quality of care in Mzimba, Malawi reported that the attitude of health workers towards clients had a negative impact on the clients' perceptions of the quality of care received (Machira & Palamuleni, 2018). As such, women preferred traditional birth attendants to skilled personnel.

The average education level of mothers is also positively associated with higher-quality antenatal care services. As previously explained, an educated woman is more knowledgeable about health education and what she should expect at antenatal care clinics during a consultation, as such, she may be able to demand higher level of quality than a woman who is not informed and does not know what to expect. Similar findings have been reported in Nigeria (Fagbamigbe & Idemudia, 2015), Nepal (Joshi et al., 2014) and Uganda (Bbaale, 2011). In line with previous studies, I also found that urban communities have better quality antenatal care services compared to rural-based clinics. A study in Nigeria also reports similar results (Fagbamigbe & Idemudia, 2015); women in urban areas were more likely to receive good-quality antenatal care compared to those in rural areas. This may be because, in most cases, urban clinics are better endowed in terms of resources and budget allocations (MoH, 2007). There is therefore a need to improve the quality of antenatal care services delivery in rural areas.

2.4.4 Strengths and limitations of the study

The findings from this study needs to be tempered by the restrictions of the methodologies employed. First, the data used in the analysis were self-reported and given by respondents about events retrospectively, which makes the collected information subjective to recall errors (I explore this further in chapter four). I made the effort to reduce this by conducting the analysis on the women with the most recent pregnancy within five years of each survey and limiting the sample size to deliveries that happened three years before and after FANC adoption. Second,

variables such as level of education, access to media through radio and/or television were measured at the time when the survey was administered and not necessarily at the time of delivery or during pregnancy as such, their impact on the outcomes of interest may not be accurate.

Despite these limitations, the major strength of the paper is that, it makes use of the nationally representative data to estimate the impact of FANC on antenatal care quality and utilisation in a low-resourced setting. Moreover, the statistical methods were able to more accurately assess the impact of FANC than the methods employed by previous studies in this research area.

2.5   Conclusion

In Malawi, the implementation of FANC has been associated with increased early access to antenatal care; however, it is also associated with the unintended consequences of underutilisation of antenatal care services and no change in the overall quality of maternal health services.

The successful implementation of FANC required investment in clinical skills, medical supplies and equipment (Lungu et al., 2011). In under-resourced health systems, this can pose a challenge. As reported in the 2010 annual monitoring report, only one of the four central hospitals and four of the 24 district hospitals in Malawi met the WHO standards for delivering FANC (MoH, 2010). Policies that target the strengthening of the capacity of health systems to implement FANC guidelines effectively would help to improve the quality and utilisation of antenatal care services. Policymakers can also explore whether further gains can be made by intensifying direct community campaigns and engagement through health surveillance assistants (HSAs), especially in rural areas where the results show relatively poor outcomes.

These recommendations apply more widely to health policy reforms in resource-constrained settings. Given that I find underutilisation when the benchmark is set at four visits, eight visits are unlikely to be feasible in settings that are struggling to comply with far less ambitious targets.

CHAPTER THREE

Optimal number of antenatal visits for positive birth outcomes in low- and middle-income countries

Abstract

Empirical evidence shows that the effective use of antenatal care improves birth outcomes and reduces the risk of neonatal mortality. Based on this suggested linear assumption, antenatal care programmes often recommend an increase in the quality of care provided (Dowswell et al., 2015). However, Malloy et al. (1992) observed that positive birth outcomes have also been reported for full-term pregnancies with fewer antenatal care visits. Similarly, other researchers have found no significant statistical differences in birth outcomes between women following a reduced-frequency model (4–6 visits) and those attending antenatal care under the more frequent model (8–16 visits). This means that there may not be a linear link between antenatal care and infant and maternal health. I explore this possibility in this chapter. I implemented a non-linear spline specification in search of turning points and thresholds in antenatal care effectiveness. Specifically, I used the nationally representative Malawi Demographic and Health Survey (DHS) data, and instrumental variable models together with highly flexible spline specifications and Wald tests to estimate breaks in the relationship between the number of antenatal care visits and the probability of low birthweight. In so doing, I established thresholds at which antenatal care stops mattering for the reduction in low birthweight. I found that the risk of low birthweight declines up until three antenatal care visits by mothers. Thereafter, the marginal effects are zero or negative, depending on the specification. Findings in this chapter suggest that the health systems of low- and middle-income countries (LMICs) may perform just as well if fewer routine visits are made with more attention to quality and that more visits should only be recommended to women who critically need them.

3.1 Background

An estimated 2.7 million newborns died in the first month of life worldwide in 2015, representing 45% of the global under-five mortality (WHO, 2016; Lawn et al., 2016). One of the principal causes of neonatal death (and a primary reason for childhood mortality and morbidity) globally is low birthweight (WHO, 2015a). In 2015, over 20 million babies were born with low birthweight worldwide; of which, 96% were from low- and middle-income countries (LMICs) (WHO, 2015a). Low birthweight, defined as birthweight of less than 2 500 g, irrespective of gestational age or any other aetiology, can be caused by preterm birth, intrauterine growth restriction, or a combination of both (Kramer, 1987). Over 80% of neonatal deaths in LMICs are due to this condition (WHO, 2012). Studies conducted by Risnes et al. (2011) and Larroque et al. (2001) found that babies with low birthweight who survive may suffer from stunting, non-communicable diseases, cardiovascular disease and neurocognitive impairment later in life. Although great progress has been made over the past decade to reduce deaths of children under five, the progress as regards reducing neonatal deaths has been slower (Koenraads et al., 2017). To reduce neonatal mortality rates and improve long-term child outcomes, targeted interventions are needed. This study analyses one such intervention, antenatal care.

Empirical evidence shows that the effective use of antenatal care improves birth outcomes and reduces the risk of neonatal mortality (Khatun & Rahman, 2008; Jewell & Triunfo 2006). Various studies have shown that the risk of low birthweight and neonatal and infant mortality increases for mothers who make fewer antenatal visits (Zhou et al., 2019; Dowswell et al., 2015; Ibrahim et al., 2012). As Dowswell et al. (2015) note, because of this linear assumption, antenatal care programmes often recommend increasing the quality of care provided, i.e. the number of antenatal visits. For example, in 2016, the World Health Organization (WHO) issued new antenatal care guidelines, which doubled the minimum number of visits for women with uncomplicated pregnancies from four to eight (WHO, 2016). However, Malloy et al. (1992) observed that positive birth outcomes have also been reported for full-term pregnancies (37–42 weeks of gestation) with fewer antenatal care visits. Moreover, Carter et al. (2016) and Majoko et al. (2007) found no significant statistical differences in perinatal outcomes between women following reduced-model (4–6 visits) and women attending antenatal care under the more frequent model (8–12 or more visits). The relationship between antenatal care and infant and maternal health may therefore not be linear. I explore this possibility in this study.

In Malawi, women with uncomplicated pregnancies are required to make a minimum of four antenatal care visits. This recommendation is based on the WHO's 2001 Focused Antenatal Care (FANC) model, which was adopted in 2003. The overall target of FANC is to reduce inadequate use of antenatal care and unnecessary visits. It emphasises the provision of goal-oriented and targeted care aimed at increasing the detection and management of complications during pregnancy (WHO, 2002a). FANC, therefore, moved away from the traditional model, which emphasised the quantity (7–16 visits) rather than quality of care.

FANC services are available in only 65% of healthcare facilities and are provided free of charge in all public hospitals in Malawi. Even though FANC is associated with improved access to antenatal care, Malawi, just like other sub-Saharan African countries, has struggled to implement the four-visit programme effectively (Mchenga et al., 2019; Lungu et al. 2011). While over 90% of pregnant women have had access to antenatal care with skilled health personnel at least once during pregnancy, only 52% of women received the required minimum four antenatal care visits (UNICEF, 2018). The reasons for the failure of the implementation of FANC has been shown to be inadequate equipment, supplies, infrastructure and training (Lungu et al., 2011; Birungi et al., 2006; Nyarko et al., 2006; Chege 2005). Furthermore, healthcare providers in LMICs have also raised concerns about the difficulty of incorporating all the FANC protocols into relatively short appointments (Maternal Health Task Force, 2014).

Given the challenges faced by LMICs in implementing FANC and achieving the less ambitious target of a minimum of four visits, the question now is, are the new antenatal care guidelines recommended or effective in low-resourced settings? Villar et al. (2001) argued that most antenatal care policies in LMICs are adopted without thorough scientific evaluation and that there is a lack of empirical evidence on the average number of visits that is likely to produce the most benefit in improving maternal and child health outcomes in settings where resources are limited. This study fills this research gap by estimating the optimal number of visits for positive birth outcomes in Malawi.

Using nationally representative Malawi Demographic and Health Survey (DHS) data, I apply instrumental variable models together with highly flexible spline specifications and Wald tests to estimate breaks in the relationship between the number of antenatal care visits and the probability of low birthweight. In so doing, I established a threshold at which an additional antenatal care visit does not have significant impact on the reduction of low birthweight. I also

explore the relationship between the number of visits and the quality of care and their independent influences on low birthweight.

The rest of the chapter is organised as follows: the next section highlights infant outcomes in sub-Saharan African countries and discusses literature on the determinants of low birthweight. The methods section, which follows, maps out the research strategy, including the dataset and the empirical method used. The subsequent sections present the results and discussion. I conclude by discussing the study's policy implications and relevance in the last section.

3.1.1   Infant outcomes in sub-Saharan African countries

A closer look at the data shows poor infant health outcomes in sub-Saharan African countries (Table 3.1). As of 2016, for example, Malawi's neonatal mortality rate is estimated at 23 per 1 000 live births, which is the same for Zambia and Zimbabwe. This figure is very high compared to high-income countries in Europe and America, which have neonatal mortality of as low as 2 per 1 000 live births and 4 per 1 000 live births respectively. Furthermore, with the prevalence rate of 12%, low birthweight in Malawi is among the highest in sub-Saharan African countries.

Table 3.1: Some key infant indicators in sub-Saharan Africa and other WHO regions

Country	Neonatal mortality rate[*]			Infant mortality rate[*]			Low birthweight
	2009	2012	2016	2009	2012	2016	2007–2016
Malawi	**30**	**24**	**23**	**69**	**46**	**42**	**12**
Zambia	29	29	23	86	56	44	11
Zimbabwe	30.5	27.8	22	58	48	38	11
Botswana	22	29	26	43	41	33	13
Kenya	27	27	23	55	49	36	8
WHO region							
Africa	36	32	28	80	63	53	14
America	9	8	4	15	13	5	8
Europe	7	6	2	12	10	3	7
Asia	31	27	28	45	39	39	30

Source: Malawi DHS data and World Bank Development Indicators (2015)
*Note: *per 1 000 live births.*

Even though low birthweight is a global phenomenon, the higher prevalence in LMICs makes it salient to understand constraints in contexts where health systems are under-resourced. Empirical evidence shows that the causes of low birthweight are complex and multifaceted, and influence either the length of gestation or the rate of intrauterine growth (Kramer, 1987). These causes include maternal age, parity, marital status, the mother's body mass index (BMI), educational attainment and social history, infant sex, prior low birthweight, malaria

39

in pregnancy, religion/ethnicity (Koenraads et al., 2017; Ngwira & Stanley, 2015; Imdad & Bhutta, 2013; Herbert et al., 2001; Kramer, 1987) and antenatal care visits (Herbert et al., 2001).

A number of studies have established a relationship between the absence of antenatal care and poor pregnancy outcomes (Dowswell et al., 2015; Mohsin, et al., 2006; Hollander, 1997; Gissler & Hemminki, 1994). This lack of care has been defined in three ways; late first access, underutilisation or inadequate number of visits, or poor quality of care (Gissler & Hemminki, 1994; Kramer, 1987). Dowswell et al. (2015) found that the risk of low birthweight and infant mortality increases if women make fewer antenatal visits (Dowswell et al., 2015). This connection may be explained by Gissler and Hemminki (1994), who highlighted two factors: either antenatal care is useful, or women with a poor diagnosis are more likely to have few visits and initiate ANC late. Additionally, Gortmaker (1979) argued that the low number of visits made by mothers might be the cause of premature deliveries. However, Malloy et al. (1992) noted that even mothers with full-term pregnancies report adverse pregnancy outcomes. The relationship between antenatal care and child and maternal outcomes may therefore not be linear, as previous researchers have generally assumed. For example, according to Carter et al. (2015), women who made many visits have reportedly undergone more labour inductions and intrapartum electronic foetal monitoring.

Now the question is: what is the lowest number of antenatal care visits that would be effective in improving birth outcomes? Such evidence in LMICs is sparse and the available evidence is mixed. For example, two studies in Zimbabwe, one in a rural (Majoko et al., 2007) and the other in an urban area (Munjanja et al., 1996) found no significant differences in low birthweight and other birth outcomes between mothers who made frequent visits (up to 14 visits) and those who made fewer visits (reduced model – four antenatal care visits). Similarly, in Nepal there were no significant differences of low birthweight incidence between mothers who made four or more visits and those receiving 1 to 3 visits (Khanal et al., 2014). In Bangladesh, Ahmed & Das (1992) also recommended three visits to reduce low birthweight.

Researchers in rural China and Brazil recommend more than four visits as being significant in reducing the incidence of low birthweight (Zhou et al., 2019; De Fonseca et al., 2014). Zhou et al. (2019) found that between five and eight visits during pregnancy are significant to prevent low birthweight in rural China. In Brazil, De Fonseca et al. (2014) found that at least seven visits were significant in reducing the incidence of low birthweight. These existing

studies, though informative, have two main weaknesses. First, they are restricted by the application of small or unrepresentative study populations (Petrou et al., 2001). Second, they are limited by the application of univariate statistical techniques, which only measure correlation and not causation between the number of visits and low birthweight. In contrast, this study uses nationally representative demographic health surveys, thus strengthening the external validity of the results to other countries with similar health system characteristics. Furthermore, this study uses quasi-experimental methods to establish a causal relationship between the number of antenatal visits and low birthweight.

3.2   Methods

3.2.1   Demographic health data

In this study, I used data from the 2015/2016 round of the nationally representative Malawi DHS. The DHS is part of the global MEASURE DHS programme, and is conducted in more than 90 LMICs. This data is publicly available at http://dhsprogram.com/data/available. The DHS gathers comprehensive health information for women of reproductive ages (15–49) and their children. A two-stage stratified cluster sample design is used which is based on the Malawi Population and Housing Census sampling frames of 2008. The response rate for the survey was 95% (NSO, 2016).

For this study, the birth recode component of the DHS was used, which contains the full birth history of all interviewed women of reproductive ages. The information in this component includes pregnancy characteristics (such as history of miscarriages/stillbirths and Caesarean deliveries) and antenatal and postnatal care for the most recent live birth within five years prior to the survey. For each child (counting multiple births as one), information is available on his or her birthweight and sex, and on his or her parents' characteristics such as age and education. Following Wehby et al. (2009), I restricted the analysis to include birthweights between 500 and 6 000 g and gestational ages between 5 and 10 months. These restrictions help to avoid potential recording errors during data collection or entry. The final total sample is therefore 13 389 babies and mothers.

3.2.2   Dependent variable: low birthweight

The main outcome variable of interest in the study is low birthweight, defined as weight at birth of less than 2 500 g (0 if birthweight $\geq$2 500 g; 1 if birthweight $<$2 500 g). Given that birthweight increases with gestational age (Gajate-Garrido, 2013; Kramer, 1987), I control for

this variable in the empirical analysis. Importantly, in the dataset, birthweight information is only available for 84.72% of all live births. If the missing values are non-random, the estimates may be biased (Mwabu, 2009). Table 1 in Appendix 1 shows significant differences in selected socio-demographic characteristics between mothers and children with reported birthweight and mothers and children without birthweight values. As such, I controlled for sample selection in the analyses using a method proposed by Olsen (1980), which I explain in the empirical analysis section.

3.2.3   Main outcome measures

The main variables of interest considered in this study are frequency of antenatal care visits and quality/content of antenatal care services.

Frequency of antenatal care visits measure

Most researchers define the extent of antenatal care as the number of visits a woman makes during pregnancy (Ibrahim et al., 2012; Gajate-Garrido, 2013; Carter et al., 2016). In this specification, they assume that the more the visits, the better the child or maternal health outcomes. However, in reality, a point of satiation beyond which antenatal care stops to matter in the influencing of outcomes may exist. In order to address this, I assumed a non-linear relationship between antenatal care and low birthweight. The antenatal care variable is therefore defined non-linearly as a set of dummies running from 0 to 8 or more.

Quality of antenatal care service measure

Further, I analysed the interaction between the facility level quality of antenatal care and the number of visits, and assessed their independent influences on low birthweight. I measured the quality of antenatal care using 10 nationally recommended antenatal care process indicators in the Malawi antenatal care guidelines (MoH, 2014) that were also captured in the DHS. The process indicators include routinely conducted diagnostics (taking blood and urine samples, measuring foetal heartbeat), physical examination (measuring blood pressure and weight), and other preventive procedures (administration of tetanus toxoid, prophylaxis, iron supplements, intestinal parasite medication and nutritional advice). Questions were asked in the DHS to assess whether women had received any one of the indicators during antenatal care visits or not, and are here interpreted as a proximate indicator of the quality of antenatal care.

42

Using the 10 quality indicators, a composite-compliance index was constructed using multiple correspondence analysis (MCA) (Greenacre & Blasius, 2006). The variables included in the composite index together with their respective categories and weights are outlined in Table 3.2. The weights are based on the first dimension of the MCA, which explains about 87.54% of the total inertia. Positive (negative) weights reflect the higher (lower) quality of care.

Table 3.2: Results for MCA for the ANC quality index

Variable	Categories	Weights
Nutrition in pregnancy	Given advice on the quality and quantity of nutrition for healthy pregnancy	0.392
	Not given any nutrition advice	-3.852
Malaria medication	Given Fansidar for the prevention of malaria in pregnancy	0.233
	Not given Fansidar	-2.293
Weight measured	Client's weight measured during visit	0.144
	Weight not measured	-6.474
Foetal heartbeat	Provider measured the foetal heartbeat during ANC	0.214
	Provider did not measure the heartbeat	-5.818
Blood pressure	Blood pressure was measured	0.575
	Blood pressure not measured	-3.305
Blood sample	Blood sample taken and tested for disease	0.304
	Blood sample never taken or tested	-4.126
Urine sample tested	Urine sample taken and tested	1.517
	Urine sample never taken or tested	-0.769
Iron tablets	Iron tablets/syrup given	0.231
	Never given iron tablets/syrup	-2.455
Albendazole	Given Albendazole for intestinal parasite	0.798
	Never given Albendazole	-0.858
Tetanus toxoid	Given at least one tetanus toxoid injection	0.178
	Never given any tetanus toxoid injection	-1.481

I then aggregated the quality index at the community level to reduce measurement error in the sample. Additionally, aggregating the quality of care measure shifts the focus from the demand side to supply side, and gives a picture of the level of services quality a facility at a community is expected to provide.

Other confounding variables that influence low birthweight

I controlled for other confounding factors that have been shown to influence low birthweight. These were categorised into five categories. The first category is maternal characteristics, which include maternal age at birth defined as a continuous variable; mother's highest education level (no education, primary, secondary and higher); marital status (married);

whether mother smokes; religion (Catholic, Muslim, Pentecostal and other); ethnicity (Chewa, Tumbuka, Yao, Tonga, Lomwe, Sena, Ngoni and other). The second group is child characteristics, which include sex of the child (boy); multiple birth (whether child is a twin); birth order (first child, 2–4th child and fifth child or more). The third group is pregnancy characteristics, which include risky pregnancy indicators (history of miscarriages and/or history of Caesarean deliveries); duration of the pregnancy before delivery (gestation age); whether pregnancy was wanted. The fourth category is household characteristics, which include household socio-economic status (poorest, poor, middle, wealthy and wealthiest); and place of residence (urban or rural). The last group is skilled antenatal care attendant (doctor/nurse). Previous studies have reported that skills acquired by antenatal care providers have a positive significant impact on the quality of service provision, which may translate into positive birth outcomes (Fagbamigbe & Idemudia, 2015; Joshi et al., 2014).

Unfortunately, most of the key clinical information, e.g. height of the mother, pre-pregnancy weight, gestational weight gain and drinking habits, were missing or non-existent in the DHS, so I could not control for them. Due to the high collinearity between the timing of the first antenatal care visit and the number of antenatal visits, I also did not control for early access in the regressions.

3.2.4   Empirical analysis

Number of antenatal care visits and low birthweight

I modelled low birthweight as a function of antenatal care utilisation and other potentially relevant characteristics that are expected to affect birthweight as follows:

$$Y_i = \beta_0 + \beta_1 Z_{di} + \beta_2 X_i + \omega_{TA} + \varepsilon_{i1} \qquad (1)$$

Where:

Y_i is 1 if child i has birthweight of less than 2 500 g and 0 otherwise;

Z is an indicator of number of antenatal care visits (defined as a set of dummy variables);

X is a vector of other factors described earlier;

ε_{i1} is a stochastic error term;

and ω_{TA} is tribal authority fixed effect, control for community differences in access to healthcare facilities.

One of the main challenges as regards estimating the impact of antenatal care is the endogeneity of antenatal care practices. Women seek antenatal care for various reasons that sometimes are difficult to observe and measure. These include the well-being of both the woman and her child, existing chronic conditions, moral values, past experiences among others (Gajate-Garrido, 2013). Rosenzweig and Schultz (1983) argued that, together, these characteristics could also determine the child's birth outcomes, which would lead to biased estimates. Researchers often apply instrumental variable (IV) methods to address endogeneity problems in antenatal care to estimate a causal and consistent effect (Rosenzweig & Schultz, 1983; Grossman & Joyce, 1990). However, when the endogenous variable is categorical or binary, as it is in this case, the use of conventional IV methods can be complicated and bivariate ordinal probit models are recommended (Buscha & Conte, 2009).

Nevertheless, Angrist (2001) argues that, even in cases where the endogenous variable is binary or categorical, the conventional IV methods still yield consistent results if the linear approximation suffices for the essentially non-linear estimation problem. Based on Angrist's argument, I applied the conventional IV methods to control for endogeneity. I estimated the first stage using the linear probability method to obtain generalised residuals from the estimated model, which were then included together with the antenatal care variable (categorised) in the birthweight equation. Following the work of Gajate-Garrido (2013) and Jensen (2000), I used rainfall shocks as the instrument. As per Gajate-Garrido's (2013) explanation, rainfall shocks do not influence the availability of health services but are related to the opportunity cost and how easy it is for a woman to access those services in a country where the majority of the roads are in poor condition.

Specifically, following the work of Gajate-Garrido (2013), I defined the instrument as standardised accrued below-average rainfall shocks during the entire pregnancy of the mother. I achieved this by linking climatic data to the locations of clusters (see Appendix 2). For every birth month, I constructed an index by subtracting the log of the monthly level of rainfall from the log of a 10-year rainfall aggregated for that period. This created variability in the instrument. Each mother and child experienced a level of below-average rainfall shock depending on the time in point the mother gave birth and not necessarily based on her location. (Gajate-Garrido, 2013). Despite the fact that climate or average rainfall can be predictable, such that mothers could possibly relocate or time their fertility accordingly, Gajate-Garrido

(2013) argues that irregular rainfall is fundamentally a random event, making below-average rainfall shocks an exogenous instrument.

Below-average rainfall shocks, as an instrument is relevant in the Malawian context because rainfall shocks have a direct impact on transportation costs and on resource and labour allocation (Atuoye et al., 2015; Gajate-Garrido, 2013). In Malawi, the quality of roads is very poor, with only 26% of the roads paved and the rest either earth or gravel (World Food Programme, 2018). Moreover, over 80% of the Malawian population resides in rural areas, where distances to the nearest healthcare facility are long and difficult to traverse during the rainy season due to poor transportation access (Varela et al., 2019). Unexpected below-average rainfall would therefore decrease the travelling times, making the roads more accessible. This would affect the mother's time availability to seek antenatal care.

In Malawi, the majority of the population depends on agriculture (Makaula et al., 2012). If there is an unexpected reduction in rainfall, less food could be produced. This could affect a mother's well-being and hence affect birthweight through nutrition (Gajate-Garrido, 2013; Bhalotra, 2010). In this study, I control for the potential direct impact of rainfall on nutrition by including the Normalised Differentiated Vegetation Index (NDVI) for each cluster, as a proxy for vegetation quality during the growing season of the mother's pregnancy (see Appendix 2). The NDVI has been shown to be highly correlated with child nutrition and health in Malawi (Johnson et al., 2013).

Gajate-Garrido (2013) argues that mother's food intake during pregnancy mostly depends on the season during which she was pregnant, as such, babies born in different seasons are likely to be exposed to different nutrients. She further explains that babies whose pregnancy period includes months of limited access to food could have relatively lower birthweights compared to babies whose pregnancy period did not. In this study, the time during which a woman was pregnant generates the variability in the rainfall instrument; to ensure that the instrument does not capture seasonality effects, I included season dummies[3] in both the antenatal care and birthweight regressions.

---

[3] I adopted the Gajate-Garrido (2013) definition of seasonality. Hence, seasonality is defined as a dummy equal to 1 for babies born between January and February and 0 otherwise, in order to account for babies whose pregnancy period included the driest months of the year in Malawi, September and October.

Henceforth, the first model I estimated was the sample-selection model to control for bias due to non-random missing values of birthweight as suggested by Olsen (1980), as follows:

$$R_i = \gamma_0 + \gamma_1\delta + \gamma_2\theta + \gamma_{3\,TA} + \varepsilon_{i2} \tag{2}$$

Where:

R_i is binary and defined as 1 if the birthweight value of child i was reported, and 0 otherwise;

δ is an instrument defined as place of delivery (1 if institution and 0 home/other) (I expect a positive relationship between institution delivery and the probability of having reported birthweight value; see Appendix 3);

θ is the vector of controls; and

ε_{i2} is a stochastic error term.

I estimated this model by ordinary least squares, obtained the predicted probabilities, $\widehat{P}$ and constructed the selection term $(\widehat{P} - 1)$, which was then included as an additional regressor in both the antenatal care and low-birthweight models.

The second model to be estimated is the first-stage antenatal care model, specified as follows:

$$Z_i = \beta_0 + \beta_1\rho + \beta_2 K + \beta_3(\widehat{P} - 1) + \omega_{TA} + \varepsilon_{i3} \tag{3}$$

Where:

Z_i is the number of antenatal care visits for child i;

ρ is below-average rainfall shocks (I expect a positive relationship between below-average rainfall shocks and the number of antenatal care visits);

K is the vector of controls that affect the number of antenatal care visits a pregnant woman makes (maternal education; household economic status; maternal age; pregnancy characteristics such as history of miscarriages, Caesarean delivery; type of residence; parity, among others);

ω_{TA} is the tribal authority (or community) fixed effects;

and ε_{i3} is the error term.

Finally, equation (1) is extended as follows:

$$Y_i = \beta_0 + \beta_1 Z_{di} + \beta_2 X_{it} + \beta_3 \varepsilon_{4i} + \beta_4 (\hat{P} - 1) + \gamma_{TA} + \varepsilon_{1i} \qquad (4)$$

Where:

Z is a set of dummies capturing the number of antenatal care visits that a pregnant woman made;

X is a vector of controls; ε_{4i} are generalised residuals from the first stage antenatal care model 3;

$(\hat{P} - 1)$ is the selection term from the sample-selection model 2;

γ_{TA} represent tribal authority (or community) fixed effects (this absorbs the common differences at local healthcare facilities); and

ε_{i1} is a stochastic error term.

Quantity vs quality of care

To analyse the relationship between the optimal number of visits and the quality of care on low birthweight, I ran simple logistic models controlling for both the optimal number of visits and the quality of care.

All the analyses were conducted using Stata Version 14.

3.3   Results

3.3.1   Descriptive statistics

The prevalence rate of low birthweight in the sample is 12%, the same as the national rate. Table 3.3 presents selected social demographic characteristics of the woman by the child's birthweight.

The average number of antenatal care visits of both women with low birthweight babies and women with babies with normal birthweight falls short of the minimum requirement of four antenatal care visits stipulated in FANC. The mean number of antenatal care visits for mothers whose babies had a low birthweight was significantly lower (but slightly low in magnitude) than those of mothers with normal-birthweight babies. Mothers without any formal education and those with only a primary-level education were more likely to have a low-birthweight baby than mothers with at least a secondary level education. Mothers from poorer households were more likely to have a low-birthweight baby than mothers from wealthy households.

Table 3.3: Social demographic characteristics by birthweight

Variables	Low birthweight		Normal weight		
	Mean	**SD**	**Mean**	**SD**	**P-value**
Number of ANC visits	3.60	1.419	3.73	1.329	0.0005
Maternal characteristics					
Age at birth	25.51	6.940	26.10	6.586	0.0005
Education level					
No education	0.12	0.329	0.11	0.308	0.0327
Primary	0.70	0.459	0.65	0.477	0.0001
Secondary	0.17	0.374	0.22	0.426	0.0000
Tertiary	0.01	0.103	0.02	0.140	0.0067
Household economic status					
Poorest	0.25	0.433	0.20	0.403	0.0000
Poor	0.23	0.422	0.21	0.406	0.0296
Middle	0.19	0.395	0.19	0.394	0.8975
Wealthy	0.17	0.378	0.19	0.395	0.0399
Wealthiest	0.15	0.359	0.20	0.401	0.0000
Pregnancy characteristics					
Gestation age	8.80	0.644	9.00	0.457	0.0000
Caesarean delivery	0.07	0.248	0.07	0.251	0.7488
Miscarriage/stillbirth	0.11	0.315	0.10	0.293	0.0299
Wanted	0.59	0.492	0.60	0.490	0.3169
Skilled ANC attendance					
Doctor/nurse	0.95	0.209	0.96	0.206	0.8039
Type of residence					
Urban	0.14	0.351	0.18	0.382	0.0005
Child characteristics					
Birth order					
1st born	0.32	0.467	0.25	0.432	0.0000
2nd–4th born	0.45	0.498	0.52	0.500	0.0000
Higher order	0.22	0.417	0.24	0.425	0.2868
Sex of the child					
Boy	0.43	0.496	0.51	0.500	0.0000
Multiple birth					
Twin	0.12	0.328	0.02	0.150	0.0000

Note: SD = standard deviation

The mean age at birth of the mothers with underweight babies (25.51 years) was significantly (but in magnitude only slightly) lower than that of the mothers of normal-weight babies (26.10 years). Mean gestation age was significantly higher for normal-weight babies (9 months) than underweight babies (8.80 months). The association between the sex of the newborn and birthweight was found to be significant. Male children were less likely to be underweight. Finally, first-born children and being part of a multiple birth increased the likelihood of low birthweight.

Distribution of number of antenatal care visits

The distribution in the number of antenatal care visits is skewed to the left, with the median number of visits being three. This means the majority of the women in Malawi do not meet the required minimum number of four antenatal care visits as recommended by the 2001 FANC model (Figure 3.1).
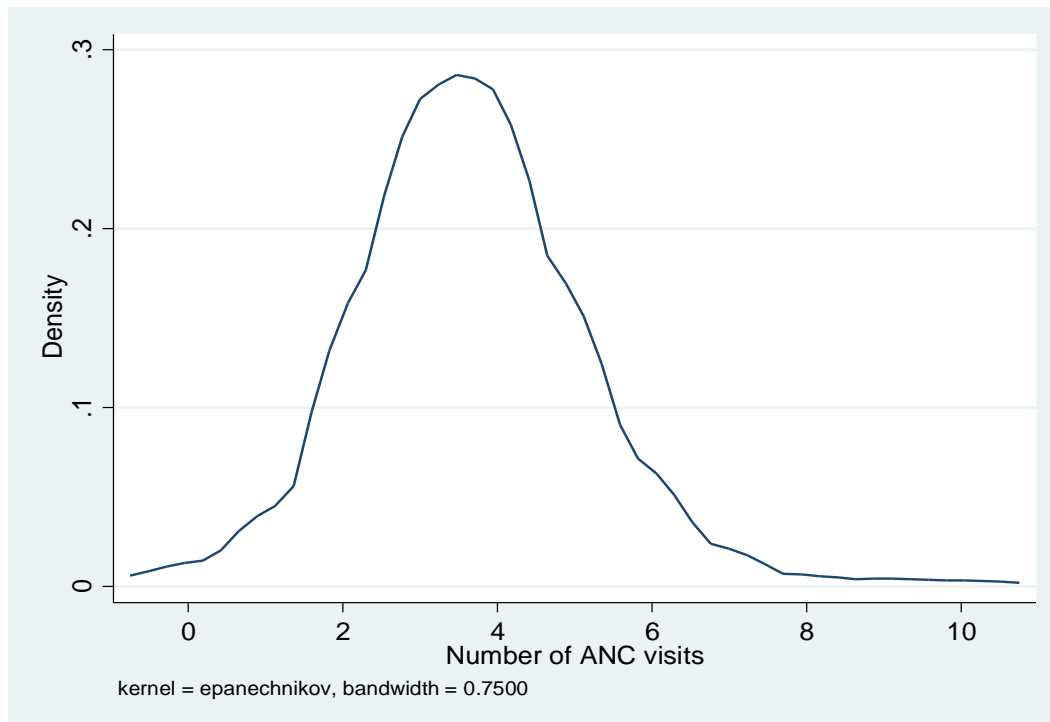
Figure 3.1: Distribution of number of antenatal care visits
Source: Own calculations from 2015/2016 DHS

Number of visits and probability of low birthweight

Figure 3.2 presents a bivariate relationship between the number of antenatal care visits and the probability of low birthweight. There is a negative relationship between the number antenatal care visits and low birthweight, i.e. the more the number of visits, the lower the probability of low birthweight. However, the negative relationship is only apparent up to four visits; beyond this level, there is little change. I further conducted regression analysis to provide causal statistical evidence on the non-linear relationship between antenatal care and low birthweight, as shown in Table 3.4.
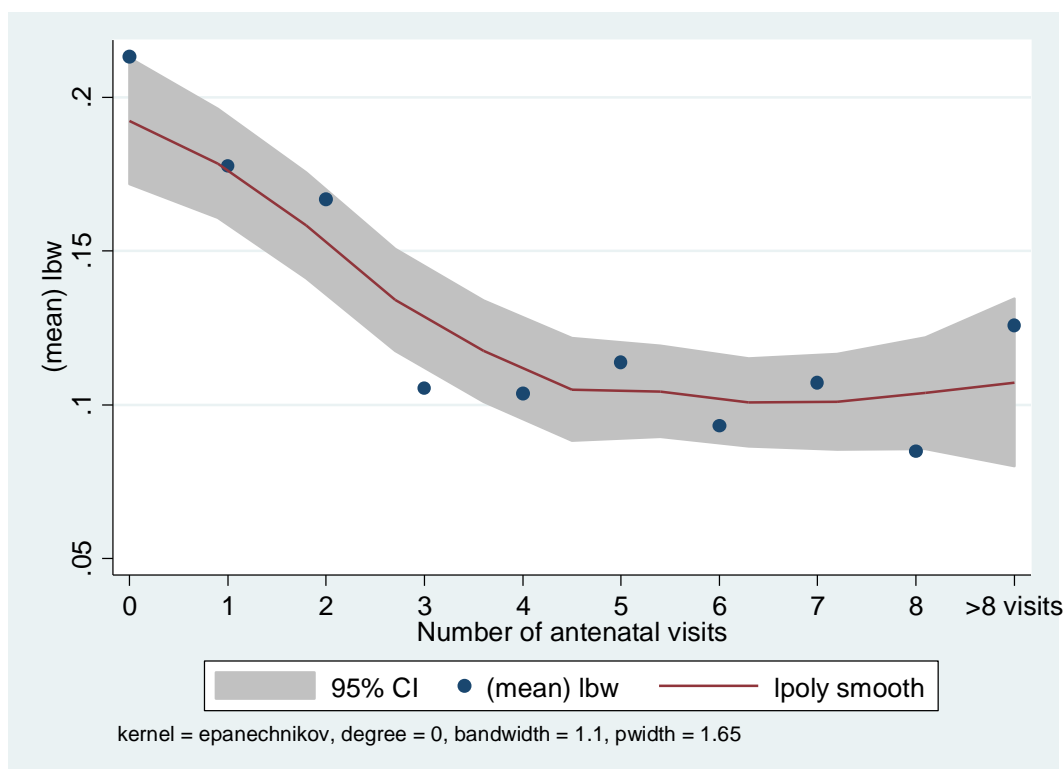
Figure 3.2: Relationship between number of ANC visits and low birthweight
Source: Own calculations from 2015/2016 DHS

3.3.2   Optimal number of ANC visits and low birthweight

Table 3.4 provides the results of Wald tests from the second stage of two-stage least square models. Results for sample selection and first-stage antenatal care regressions are reported in Appendix 4. The tests examine structural breaks in the relationship between antenatal care visits and low birthweight. The first part of the hypothesis tests whether each additional visit adds value over and above the previous visit; I also tested for second differences, to assess whether the profile accelerates or decelerates significantly at each point. In this way, I detected significant breaks in the relationship between antenatal care and low birthweight similar to other estimation contexts (McKenzie, 2006). Because many of the specifications contain generated regressors from selection and first-stage IV equations, I bootstrapped the specifications by 100 replications (Wooldridge, 2002).

I estimated the following model specifications:

- ordinary least squares (OLS);
- OLS controlling for sample selection;
- OLS controlled for both sample selection and community fixed effects;
- two-stage least squares (2SLS) controlling for sample selection; and

51

- 2SLS with sample selection and community fixed effects.

Even though the instrument is strong in the IV regressions (with F-statistics of greater than 10), the Hausman test in the respective regressions is insignificant, showing that the IV models are not statistically different from the OLS models. I therefore continue to analyse and interpret the models without IVs. On the other hand, the sample-selection variable is significant in all models, confirming the descriptive statistics that the bias due to non-random missing of birthweight data needs to be controlled.

For the individual visits, OLS results show that 3–8 antenatal care visits deliver significantly better birthweight outcomes compared to no visits. In the first differences, however, three visits add significant value over two visits. This finding is strengthened by the analysis of second differences, showing a significant acceleration in the reduction of low birthweight at three visits across all specifications. Therefore, a clear structural break emerges at three visits, beyond which additional antenatal care does not appear to improve low birthweight.

Table 3.4: Wald test results for birthweight models

	OLS 1	OLS 2	OLS 3	IV1	IV2
Null hypothesis	**Basic**	**SS**	**SS_FE**	**SS**	**SS_FE**
Relative to no visits					
1 ANC visit=0	-0.036	-0.013	0.010	-0.034	-0.035
	(0.059)	(0.060)	(0.059)	(0.056)	(0.056)
2 ANC visit=0	-0.046	-0.027	-0.009	-0.029	-0.027
	(0.051)	(0.049)	(0.048)	(0.047)	(0.046)
3 ANC visits=0	-0.108**	-0.086*	-0.070	-0.072	-0.066
	(0.048)	(0.047)	(0.046)	(0.050)	(0.048)
4 ANC visit=0	-0.110**	-0.086*	-0.073	-0.061	-0.052
	(0.048)	(0.047)	(0.046)	(0.058)	(0.053)
5 ANC visit=0	-0.099**	-0.074	-0.058	-0.040	-0.028
	(0.050)	(0.048)	(0.046)	(0.068)	(0.060)
6 ANC visit=0	-0.120**	-0.094*	-0.079*	-0.047	-0.032
	(0.050)	(0.049)	(0.048)	(0.078)	(0.066)
7 ANC visit=0	0.106*	-0.078	-0.065	-0.018	-0.0003
	(0.056)	(0.055)	(0.055)	(0.094)	(0.079)
8 ANC visit=0	-0.128**	-0.101	-0.108*	-0.033	-0.012
	(0.063)	(0.063)	(0.063)	(0.106)	(0.087)
>8 ANC visit=0	-0.087	-0.058	-0.036	-0.017	-0.041
	(0.060)	(0.059)	(0.059)	(0.116)	(0.095)

	OLS 1	OLS 2	OLS 3	IV1	IV2
Null hypothesis	**Basic**	**SS**	**SS_FE**	**SS**	**SS_FE**
Marginal change relative to one less visit					
2.ANC visit - 1.ANC visit=0	-0.011	-0.014	-0.019	0.005	0.009
	(0.037)	(0.038)	(0.040)	(0.039)	(0.039)
3.ANC visit - 2.ANC visit=0	-0.062***	-0.059***	-0.061***	-0.042**	-0.039**
	(0.016)	(0.016)	(0.016)	(0.020)	(0.018)
4.ANC visit - 3.ANC visit=0	-0.002	-0.001	-0.003	0.010	0.014
	(0.009)	(0.009)	(0.009)	(0.015)	(0.013)
5.ANC visit - 4.ANC visit=0	0.010	0.012	0.015	0.021	0.024
	(0.012)	(0.012)	(0.012)	(0.018)	(0.015)
6.ANC visit - 5.ANC visit=0	-0.021	-0.020	-0.021	-0.007	-0.004
	(0.018)	(0.017)	(0.018)	(0.022)	(0.020)
7.ANC visit - 6.ANC visit=0	0.014	0.015	0.014	0.029	0.032
	(0.032)	(0.034)	(0.032)	(0.036)	(0.035)
8.ANC visit - 7.ANC visit=0	-0.022	-0.023	-0.043	-0.015	-0.011
	(0.052)	(0.051)	(0.053)	(0.051)	(0.050)
9.ANC visit - 8.ANC visit=0	0.041	0.042	0.073	0.050	0.053
	(0.055)	(0.055)	(0.057)	(0.056)	(0.056)
Structural breaks					
3.ANC-2.ANC=2.ANC-1.ANC	-0.051	-0.045	-0.042	-0.047	-0.048
	(0.044)	(0.046)	(0.048)	(0.045)	(0.045)
4.ANC- 3.ANC=3.ANC-2.ANC	0.060***	0.058***	0.058***	0.053***	0.053***
	(0.019)	(0.020)	(0.020)	(0.019)	(0.019)
5.ANC-4.ANC =4.ANC-3.ANC	0.012	0.013	0.017	0.011	0.011
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
6.ANC-5.ANC= 5.ANC-4.ANC	-0.031	-0.032	-0.035	-0.028	-0.028
	(0.026)	(0.025)	(0.025)	(0.026)	(0.026)
7.ANC-6.ANC=6.ANC-5.ANC	0.035	0.035	0.034	0.036	0.036
	(0.043)	(0.043)	(0.041)	(0.043)	(0.043)
8.ANC-7.ANC=7.ANC- 6.ANC	-0.037	-0.038	-0.056	-0.043	-0.043
	(0.075)	(0.074)	(0.074)	(0.073)	(0.073)
9.ANC–8.ANC=8.ANC-7.ANC	0.064	0.065	0.115	0.065	0.064
	(0.096)	(0.095)	(0.100)	(0.092)	(0.092)
Sample-selection variable	-	-0.266***	-0.276***	-0.237***	-0.248***
		(0.056)	(0.054)	(0.056)	(0.057)
Seasonality	No	No	No	Yes	Yes
Gestation	Yes	Yes	Yes	Yes	Yes
Child characteristics	Yes	Yes	Yes	Yes	Yes
Risky pregnancy indicators	Yes	Yes	Yes	Yes	Yes
Skilled ANC attendance	Yes	Yes	Yes	Yes	Yes
Mother characteristics	Yes	Yes	Yes	Yes	Yes
Household characteristics	Yes	Yes	Yes	Yes	Yes
Cultural and religious factors	Yes	Yes	Yes	Yes	Yes
Community fixed effects	No	No	Yes	No	Yes
Instrumental variable from the first-stage below-average rain	-	-	-	0.037***	0.029**
				(0.009)	(0.012)
Hausman test	-	-	-	-0.009	-0.015
				(0.013)	(0.014)
F-statistics			-	26.79	11.43
Number of observations	11 634	11 593	11 593	11 593	11 593

*Notes: *, ** and *** indicate significant effects at P<0.1, P<0.05 and P<0.01, respectively.*

3.3.3   Quantity vs quality of care

In section 3.3.2 of this chapter, I established that at least three antenatal care visits are associated with a reduction in low birthweight. In this section, I explore the relationship between the optimal number of visits and quality of care on low birthweight. Previously I hypothesised that offering many and frequent visits to mothers may compromise the quality of care in LMICs. The bivariate results in Figure 3.3, however, shows that in communities where quality of care is high, women are more likely to make more visits.
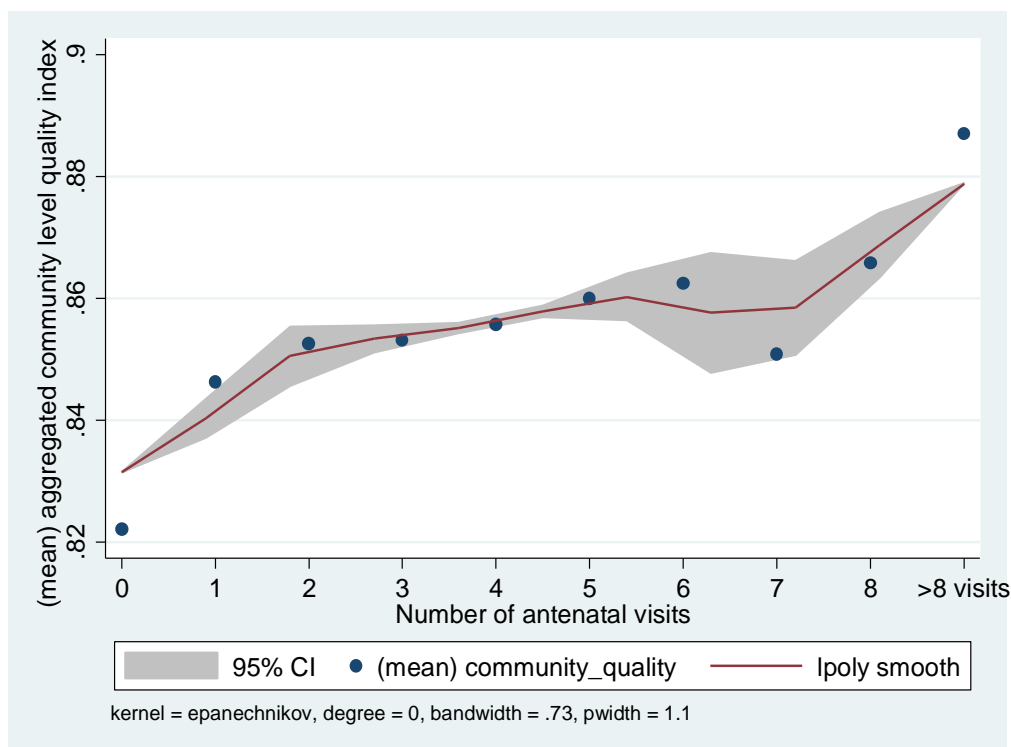
Figure 3.3: Number of ANC visits and quality of care
Source: Own calculations from 2015/2016 DHS and 2013/2014 SPA

As regards the relationship between quality of care and probability of low birthweight, the results in Figure 3.4 show no direct relationship between the two. I further explored the relationship between the quantity and quality of antenatal care and low birthweight using logistic regressions, reported in Table 3.5. I ran five regression models, as follows, model (1) with just a community-level quality of care index; model (2), which controls for both quality of care and the structural break of three visits; model (3), which controls for other confounding factors; model (4), which assesses the combined effect of the quality and quality of care by

54

interacting the two variables. Finally, model (5), besides the interaction variables, also controls for other confounding factors that influence low birthweight.
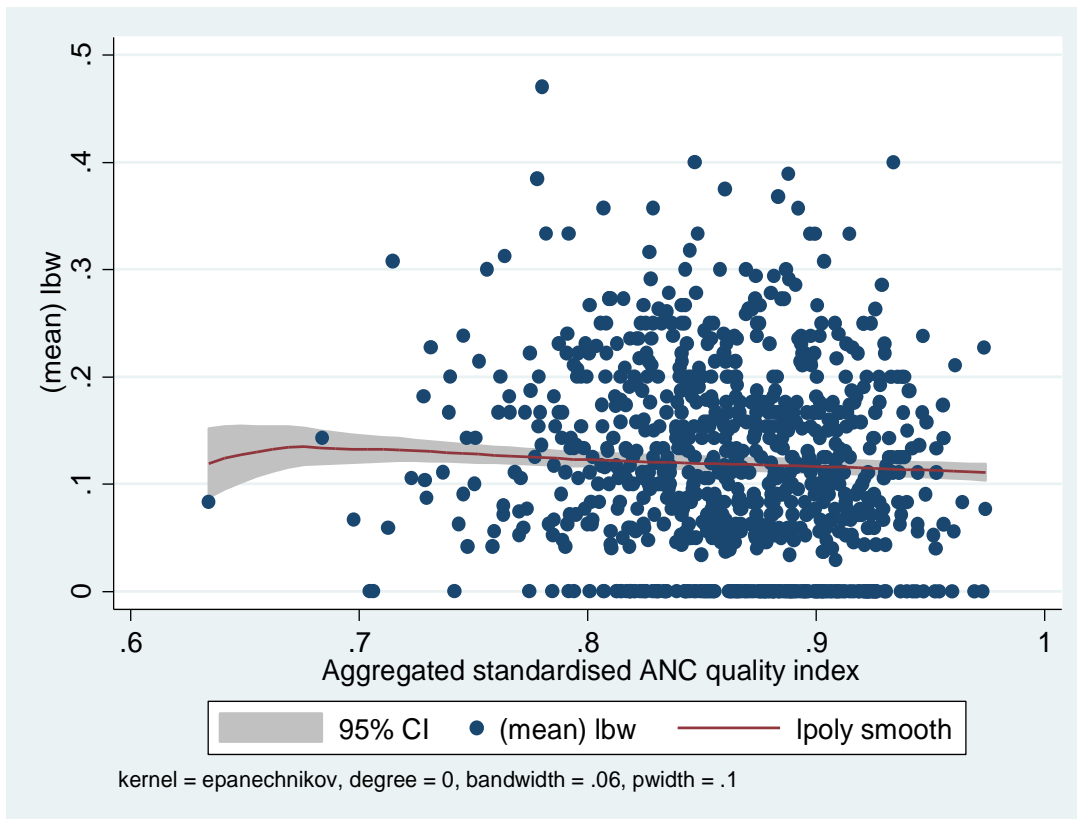
Figure 3.4: Quality of ANC and low birthweight
Source: Own calculations from the 2015/2016 DHS

There is a significant and negative relationship between the quality of antenatal care and low birthweight in the first model. However, the significance disappears when I controlled for number of visits and other confounding variables. Furthermore, not only are the interaction variables between quality and the number of visits not significant but variables capturing the structural break also become insignificant in models 4 and 5. However, due to data limitation, I was not able to disentangle the combined effect of number of antenatal care visits and quality of care on low birthweight, there is need to further explore this relationship using qualitative methods.

Other factors that reduce the likelihood of low birthweight include a longer gestation period; being a male child; later births (birth order); maternal age, although the relationship is not linear; a higher level of education; and higher socio-economic status.

Table 3.5: Marginal effects for the impact of quality and quantity of ANC on low birthweight

Low birthweight	(1)	(2)	(3)	(4)	(5)
Community quality of care index	-0.135* (0.070)	-0.110 (0.070)	-0.079 (0.069)	0.102 (0.350)	-0.048 (0.339)
Spline ANC visits					
Up to 3 ANC visits		-0.034*** (0.006)	-0.033*** (0.007)	0.007 (0.107)	-0.040 (0.105)
4+ ANC visits		-0.0004 (0.003)	0.003 (0.003)	0.078 (0.052)	0.059 (0.049)
Quality vs visits					
ANC community quality*3 ANC visits				-0.049 (0.126)	0.008 (0.123)
ANC community quality*4 ANC visits				-0.091 (0.060)	-0.065 (0.057)
Pregnancy factors					
Gestation			-0.072*** (0.006)		-0.072*** (0.006)
Wanted pregnancy			-0.006 (0.007)		-0.006 (0.007)
Child factors					
Sex of the child-boy			-0.022*** (0.007)		-0.022*** (0.007)
Child is a twin			0.279** 0.041)		0.279*** (0.041)
Birth order (first-born) 2–4th			-0.047*** (0.012)		-0.047*** ().012)
>4th			-0.057*** (0.013)		-0.057*** (0.013)
Pregnancy risk factors					
History of miscarriage			0.034** (0.014)		0.034** (0.014)
History of Caesarean			-0.006 (0.013)		-0.006 (0.013)
Age at birth			-0.008* (0.004)		-0.008* (0.004)
Age*Age			0.0002** (0.000)		0.0002** (0.000)
Maternal education (Ref – No educ) Primary			-0.006 (0.012)		-0.006 (0.012)
Secondary			-0.028** (0.013)		-0.029** (0.013)
Tertiary			-0.051** (0.020)		-0.049** (0.021)

Low birthweight	(1)	(2)	(3)	(4)	(5)
Socio-economic status					
(Ref – poorest)					
Poor			-0.003		-0.003
			(0.010)		(0.010)
Middle			-0.009		-0.009
			(0.009)		(0.009)
Wealthy			-0.014		-0.014
			(0.010)		(0.010)
Wealthiest			-0.026**		-0.026**
			(0.012)		(0.012)
Other controls					
Skilled ANC attendant	No	No	Yes	No	Yes
Ethnic background	No	No	Yes	No	Yes
Smoking	No	No	Yes	No	Yes
Urban location	No	No	Yes	No	Yes
Nutrition proxy – NDVI	No	No	Yes	No	Yes
N	11 634	11 634	11 593	11 634	11 593

*Notes: *, ** and *** indicate significant effects at P<0.1, P<0.05 and P<0.01, respectively*

3.4   Discussion

In the present study, I estimated the relative number of antenatal visits that is likely to make the most difference in reducing the probability of low birthweight in a low-resourced setting. I also assessed both the individual and the combined impact of the optimal number of visits and quality of care on low birthweight.

First, I found a structural break at three antenatal care visits, beyond which the marginal benefit of an additional visit is zero (I found the same results using bivariate ordinal probit models). Similar findings have been reported in Bangladesh (Khutan & Rahman, 2008; Ahmed and Das, 1992) and Nepal (Khanal et al., 2014) where a minimum of three visits was recommended to be significant in reducing the incidence of low birthweight. The results also confirm findings from Zimbabwe where no significant differences in low birthweight and other birth outcomes between women who received the traditional model of antenatal care (up to 14 visits) and those who received the reduced model (minimum of four visits) were found (Mujaja, 1996; Majoko et al., 2007).

The strength of this study, however, lies in the methodology used. Unlike the previous studies, which used dose-response analysis or qualitative methods to show changes in birth outcomes across different models of antenatal care, this study used quasi-experimental methods. I have therefore been able to show a causal relationship between the number of visits and low

birthweight. Furthermore, I found a non-linear effect of antenatal care on birthweight that does not disappear or weaken after controlling for gestation age, as other linear studies have reported (Gajate-Garrido, 2013; Abdal Qader et al., 2012). Overall, based on the findings, a reduced model of antenatal care may be the better option for Malawi.

In Malawi, doubling the minimum number of visits from four to eight, as the 2016 WHO guidelines suggest, is also an economic question. For example, a somewhat dated study found that the unit societal cost (household and facility) of one antenatal care visit ranges between $3.2 and $5.8, depending on the level of service provision and type of provider in Malawi (Levin et al., 2003). Even though the cost may seem to be low in comparison to other LMICs (unit cost of antenatal care in Tanzania ranges from $2.78 to $59.48), in a country where over half the population lives below $1.90 per day, the current cost is high. According to a recent study in Rwanda, the full utilisation and implementation of the eight-visit model would increase costs by 18% compared to the four-visit model (Hitimana et al., 2018). In Malawi, the increase may be more, given the inefficiencies that the health system is currently facing (MoH, 2014a).

Second, I also hypothesised that increasing the number of antenatal care visits beyond a certain threshold is likely to compromise the quality of care. Results found by Benova et al. (2018) confirm this hypothesis. The researchers observed low quality of antenatal care in selected LMICs even among women with a reported minimum number of four visits. In this study, however, contrary to the hypothesis and the findings by Benova et al. (2018), bivariate analysis shows that women are more likely to make more visits in facilities where the quality of care is high, with significant increases starting from three visits up to eight visits. On the other hand, the regressions show no direct impact of quality of care on low birthweight when I controlled for the number of visits and other confounding variables. It could be that there are mechanisms between the number of visits and the quality of care that are driving the threshold effect on low birthweight. However, in the context of this study, it is not clear which mechanisms they might be.

Nevertheless, given that critical assessment during antenatal care often happens at the beginning and in the middle of pregnancy – usually by the third antenatal care visit (WHO, 2016; WHO, 2002a) – three visits may be enough to provide the level of quality necessary to improve birthweight. Beyond the fourth visit, not much is done except for the usual

assessments to monitor if there are no complications prior to delivery. The main justification for the revision of the minimum number of visits from four to eight by the WHO is that the reduced model of care was associated with an increased risk of perinatal mortality and missed complications compared to models with a minimum number of eight visits (Dowswell et al., 2015). I acknowledge that, indeed, more visits could possibly have a significant impact on other clinical outcomes, which I could not show here. However, in settings where resources are limited, investing in fewer but higher-quality visits can prove effective and more visits can be scheduled only for women with the potential for complications, just as the 2001 guidelines suggest (WHO, 2001).

Meanwhile, I did find socio-economic factors to have a significant impact on low birthweight. For example, the mother's education was significantly associated with a reduction in the probability of low birthweight. Similarly, women from wealthier households were more likely to have normal-birthweight babies than women from poorer households. Other researchers have also reported similar findings (Khatum & Rahman 2008; Kramer 1987). Moreover, Filmer and Pritchet (1999) found that variations in under-five child mortality were mostly explained by the socio-economic factors, for example the level of female education, distribution of income, religion and ethnicity.

The major strengths of this chapter are based not only on the nationally representative data used but also on the methodologies applied to establish a causal relationship between the actual number of visits and birth outcomes. This study departed from the linearity assumption by using flexible non-linear specifications and controlling for endogeneity and sample selection. I established a threshold beyond which marginal improvements in low birthweight disappear. However, this kind of analysis is not without weaknesses. The data used in the analysis were self-reported and the survey was conducted retrospectively, making the collected information prone to errors in recall (I further explore this issue in the next chapter). An effort was made to minimise recall bias by conducting the analysis on women with the most recent pregnancy within five years prior to the survey. Furthermore, given small samples of high-frequency antenatal care visits in the sample (most of the women made an average of only three visits), the results could be overestimating the significance of three visits. Future work should therefore explore the use of multi-country-level data and other birth outcomes to increase the variation in the number of visits and the external validity of the results.

3.5   Conclusion

This chapter estimated the optimal number of antenatal care visits that is associated with a reduction in low birthweight deliveries in Malawi. I did this by applying conventional instrumental variable models in combination with Wald tests and spline regressions. I found that up to three visits significantly reduced the probability of low birthweight in Malawi. Beyond that, no marginal benefits arise. Furthermore, even though there was a significant relationship between the number of antenatal care visits and the quality of care, the results show no direct impact between the quality of care and low birthweight. Given resource and cost constraints, it is concerning that there does not appear to be a strong rationale for doubling the minimum number of visits to eight in this study's context, as suggested in the 2016 WHO antenatal care guidelines. However, I acknowledge that doubling the minimum number of visits to eight may have an impact on other maternal and child outcomes (which unfortunately I was unable to show here due to data limitations). I therefore recommend that further research should be conducted to assess the effectiveness of the 2016 antenatal care guidelines in improving birth outcomes in low-resourced settings before the guidelines are fully implemented.

Dowswell et al (2015) argued that a probable consequence of the four-visit model could be more missed diagnoses or complications. In a retrospective study of births in Singapore, Sen et al. (1991) found that a schedule of three visits per pregnancy is likely to miss 16% of major pregnancy complications. However, other studies did not find this to be true and reported otherwise (Carter et al., 2016; Majoko et al., 2007; Jewell et al., 2000). In resource-constrained environments, such evidence is necessary for efficient and effective policy implementation. While the issues raised above are relevant across the world, they are especially relevant for LMICs whose health systems are severely overburdened and under-resourced.

CHAPTER FOUR

Measuring antenatal care quality: Comparing estimates from household and facility survey data in Malawi

Abstract

Relatively high antenatal care (ANC) coverage continues to coexist with poor maternal and child health outcomes in low- and middle-income countries (LMICs). This weak relationship between the use of antenatal care and maternal and newborn survival has moved the focus from the quantity of care to the quality of care provided. Accurate information on the quality of care is therefore important to improve maternal healthcare interventions and, hence, outcomes. In LMICs, estimates of quality of care are largely obtained from exit interviews and household retrospective surveys, which are based on self-reports. However, self-reported data has well-documented weaknesses. Among other things, these include recall errors and courtesy bias, which affect reporting accuracy and limit the utility of survey results for programme improvement. Despite the well-documented weaknesses of self-reported data, there is limited empirical research validating the accuracy of women's self-reports on the quality of antenatal care in LMICs. This chapter seeks to fill that research gap. Using direct client observations at a facility as a gold standard, I compared these observations to women's self-reports in the exit interviews after the antenatal care consultation and retrospective self-reports a year after delivery as captured in the Malawi Demographic and Health Survey (DHS). Specifically, I tested the sensitivity, specificity and receiver-operating curves (ROCs) of quality indicators of antenatal care as captured in the three survey methods. The results suggest that women overestimate the quality of care received in comparison to the actual observed quality of care due to lack of information on the content of care and also not understanding the survey questions. The findings from this chapter suggest that, in measuring the quality of antenatal care for monitoring, it is important to compare women's self-reports with direct observations or facility data in order to get accurate quality estimates.

4.1   Background

Relatively high antenatal care coverage continues to coexist with poor maternal and child health outcomes in low- and middle-income countries (LMICs) (Alkema et al., 2016; Kambala et al., 2015; Kyei et al., 2012). In Malawi, for example, even though 95% of women attend antenatal care at least once, in 2016, the country registered one of the highest maternal mortality rates in the sub-Saharan region: 497 deaths per 100 000 live births and neonatal mortality of 27 deaths per 1 000 live births (NSO, 2016). This overlap between the increased use of antenatal care and poor survival rate of maternal and newborn has moved the focus from the quantity of care to the content and overall quality of care provided (Benova et al., 2018; Kyei et al., 2012). In light of this paradigm shift, Benova et al. (2018) argue that global and national progress towards quality universal coverage can only be monitored if appropriate, valid, reliable and complementary tools and data sources are used when capturing indicators of quality.

In both high-income countries and LMICs, a number of data collection methods has been used to monitor the quality of services provision. These include direct clinical observations of the delivery of the services; medical record reviews; interviews and focus groups with patients after receiving care; provider interviews; and retrospective self-reports (Bhattacharya et al., 2019; Kyei et al., 2015; Hodgins, 2013; Hermida et al., 1999). In high-income countries, the most common method used to assess service quality is the review of clinical medical records and other written records (Donabedian, 1980; Hermida et al., 1999). These allow retrospective assessment of routine provider performance, do not require a lot of resources, and are relatively easy to evaluate (Onishi et al., 2011; Franco et al., 2002). However, in LMICs, medical records are rarely used due to inadequate, inconsistent or even missing records, predominantly at government owned facilities (Choi 2018; Hermida et.al 1997). As such, quality of care estimates are usually derived from the following three data collection methods:

1. direct observations of care;
2. client exit interviews; and
3. retrospective self-reports based on household surveys such as the Demographic and Health Survey (DHS) and Multiple Indicator Cluster Survey (MICS).

Various researchers have thoroughly assessed the highlighted data collection methods as regards their strengths, flaws, viability and types of information each method can collect (Choi

2018; Onishi et al., 2011; Masatu et al., 2006; Franco et al., 2002; Hermida et al., 1999). Despite of this, limited empirical research has been conducted to directly compare the different methods with one another to examine their validity in measuring the quality of antenatal care. While some authors have assessed the validity of women's reports on maternal newborn and child health outcomes and healthcare coverage in the postnatal period in LMICs (Olatunde et al., 2018; Blanc et al., 2016; Liu et al., 2013; Stanton et al., 2013), this study focused exclusively on the level of service quality in the antenatal period. Direct observations were considered as a gold standard and were compared with;

- exit interviews with pregnant women after antenatal care consultation; and
- retrospective self-reports on the provision of antenatal care services as captured in the Malawi DHS.

Specifically, I tested the sensitivity, specificity and receiver-operating curves (ROCs) values of antenatal care quality indicators as captured in the three methods mentioned above.

The rest of the chapter is organised as follows: I begin by discussing the potential biases of different quality assessment methods before delving into the literature review, methodology, analysis, results and discussion.

4.2 Potential biases of the three data collection methods for estimating quality of care in LMICs

Figure 4.1 is a framework describing the potential weaknesses and biases of the three quality assessment methods used to collect the quality of care data that was analysed in this study. Clinical direct observation, which involves recordings of the provider's actions during a consultation by an independent observer, is mostly considered a gold standard (Blanc et al., 2016). Franco et al. (2002) argued that information derived from direct observation, when the independent observer simultaneously records the providers actions using structured checklist to assess whether he/she is following a set of guidelines, has the potential to provide one of the most complete and reliable pictures of what providers do. The checklist used by the observer has a list of interventions that a provider is expected to perform during a consultation based on the antenatal care guidelines. For example, the observer checks whether the provider has given or prescribed iron or folic acid, and indicates '1' if he or she observed the provider providing the service, and '0' otherwise.

The disadvantage of direct observations, however, is that they only measure quality of care at one point in time. Given that not all interventions of antenatal care are administered at every visit, direct observations may underestimate the quality level of care. Moreover, direct observations are only conducted in facilities where providers report to have antenatal care services, leading to sample-selection bias. This may cause an upward or downward bias in the quality estimates based on the observed facility characteristics. One of the well-documented weakness of direct observation is that providers modify their behaviour when they know they are being observed, a phenomenon called the Hawthorne effect (Mayo, 1933; Haessler, 2014; Chen et al., 2013), hence observed performance might not represent routine performance, leading to upward bias in quality. Lastly, direct observations are also considered fairly invasive and costly, as they require a trained observer to follow a patient through the clinic.

Client exit interviews, on the other hand, are less intrusive and are cost effective. They may capture routine performance if they are conducted without the provider knowledge, although trained interviewers are still required to speak to patients or their parents or guardians (Hutchinson et al., 2011; Gilchrist et al., 2004). Just like direct observation, client exit interviews have their own weaknesses. First, the sample in the client interviews is limited to women who visited the facility and gave consent to be interviewed (Franco et al., 2002), leading to sample-selection bias if selection is non-random. Second, a structured questionnaire is used to assess the level of satisfaction or the quality of care. This requires some level of understanding and knowledge of the content of care by the client and the ability to link the questions to what the provider was doing and discussed. In the context of this study, examples of questions in the client exit interview include whether the provider prescribed or gave iron or folic acid during in the current visit only, during the previous visit only, or during both visits. As regards counselling, clients are asked whether the provider talked about the side effects of iron tablets in the current visit, the previous visit, or both. The two highlighted limitations can either lead to understating or overstating the level of satisfaction on the services received.

The main shortcoming of this approach as noted by Hameed et al. (2018), however, is the likelihood of 'courtesy bias' in the responses of the clients. Avis et al. (1997) explained that courtesy bias happens when a client may be hesitant to share negative experiences because the service provider is close by or within earshot This high level of reported satisfaction with services may limit the utility of survey results for programme improvement (Lindelow, 2003).
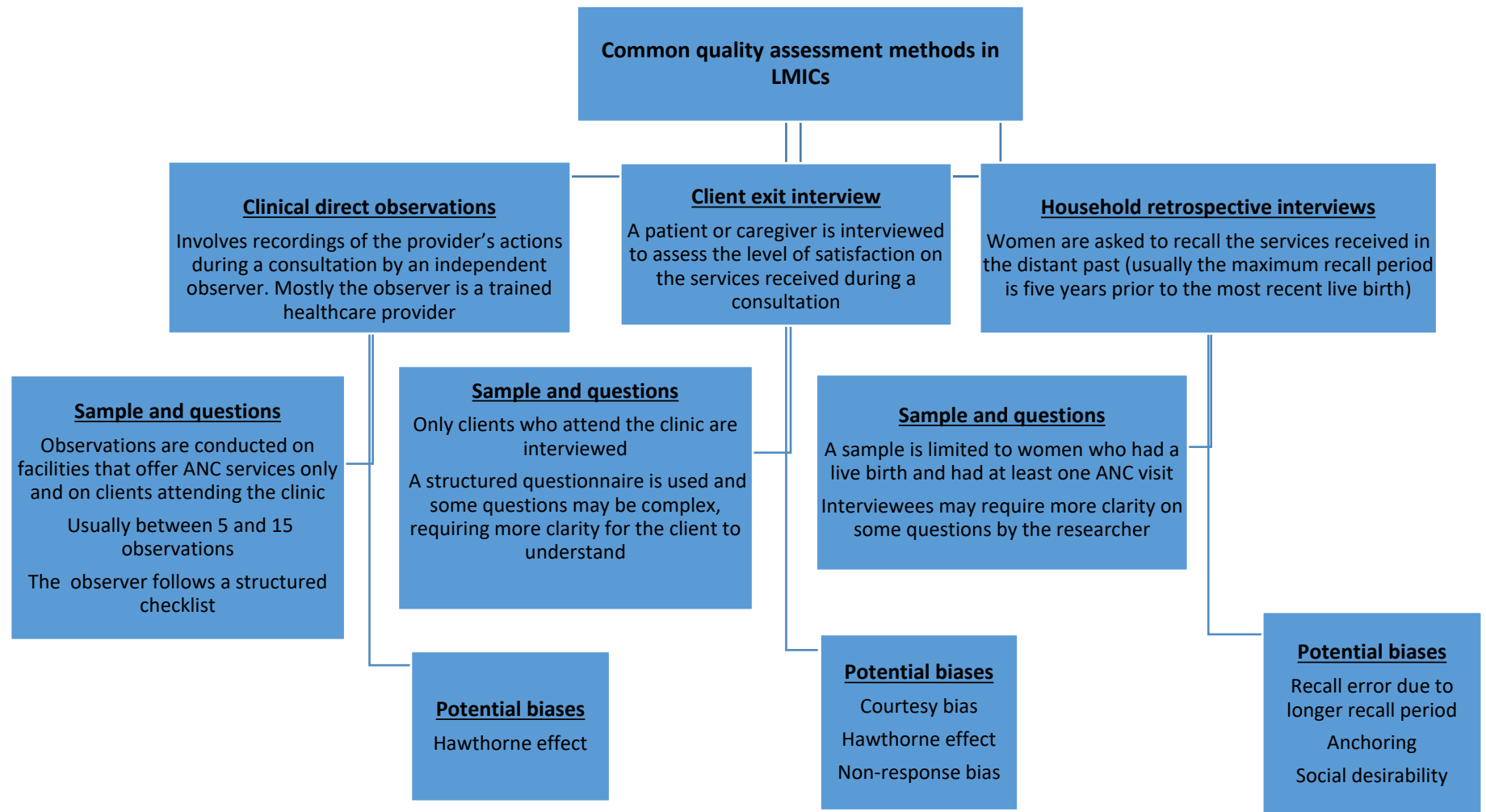
Figure 4.1: Framework describing different assessment methods of collecting quality of care data in LMICs
Source: Own work

Similar to client exit interviews, retrospective household surveys can also be implemented on a larger scale and are cost-effective (Choi, 2018). The Demographic and Household Surveys (DHS) and Multiple Indicator Cluster Surveys (MICS) are well-known surveys that conduct retrospective interviews on quality of care in LMICs. These surveys collect detailed data on the content of maternal and child health services received from women who had a live birth in the five years prior to the survey. A structured questionnaire is also administered, asking women whether at any point in the pregnancy, a certain intervention was offered at least once during antenatal care. Unlike direct observation of a client at a facility and client interviews, household surveys measure the quality of service provision over the entire pregnancy period. As such, there is a possibility that some interventions that are rated very high in household interviews were not a requirement for a particular woman's condition based on the gestation of her pregnancy in the direct observation or client exit interview.

Even though the wide availability of the surveys is an advantage for global-level monitoring in low-resourced settings, just like client exit interviews, the usefulness of these data hinges on a number of factors. It requires that women are able to clearly understand the questions being asked, are willing to participate and share their answers honestly, and have knowledge about the components of care they received. Furthermore, retrospective surveys have a longer recall period (up to 5 years), therefore their reliability also depends on whether the woman is able to provide a consistent recollection of what happened in the past, and that a woman's current characteristics or conditions do not significantly influence her ability to accurately remember what happened (anchoring) (Von Fintel & Posel 2015). For example, if a child had satisfactory birthweight, and is growing up well, mothers may be more likely to recall quality of care in a positive light, leading to upward bias in quality measurement.

Given that comparable data is often limited in LMICs, the basic question that often arises is how to interpret results from various data sources (Choi, 2018). This study assesses the congruency of women's reports in the client exit interviews and retrospective household surveys with direct client observation at a facility.

4.2.1 Validation studies on maternal health services in LMICs

Very few studies have evaluated the reliability and validity of self-reports on the provision of maternal healthcare services in LMICs. The existing LMIC studies report low to moderate

validity across the studied indicators of maternal health services. For example, in their study of perinatal depression in Nigeria, Ayinde et al. (2018) found that most of the pregnant women and those who had just delivered rated the service provided in the facilities as being of good quality and reported higher satisfaction levels with the care received, notwithstanding the objective evidence of poor quality of care of perinatal depression. In Kenya, Blanc et al. (2016) compared observation records and women's reports in the exit interviews on skilled birth attendants. They found that women's reports of skilled attendant met the validation criteria only for population level indicators and the reporting accuracy was very low where there was low skilled attendant coverage (Blanc et al. 2016). A study in Mozambique, analysing the consistency of women's self-reports on key maternal and child interventions during the peripartum period, found that women were able to consistently report on selected peripartum care aspects but not all of them (Stanton et al., 2013). Similarly, researchers in rural China found the overall validity of self-reported coverage moderate across selected maternal and child health indicators (Liu et al., 2013).

The weakness in these studies is that they use qualitative and primary data, limiting the sample to few facilities and women; hence, their external validity is limited. One exception, however, is the study by Choi (2018), which examined the consistency of women's reports of side effects counselling in family planning using the nationally representative household Demographic and Health Surveys and facility census surveys. Choi used two definitions of counselling to determine whether the health workers explained common side effects to the women (minimum counselling), and whether health workers furthermore also explained to the women what they needed to do in case of side effects (expanded counselling). The researcher found that the validity of women's reports depended on the definition of counselling used. Women's reports were relatively similar to the observation-based estimates when the minimum definition of counselling was applied, but inconsistent when the expanded definition was applied.

This study contributes to this body of scarce data validation literature in LMICs and the ongoing debate on the need for accurate data to effectively measure quality of care.

4.3   Data and methods

4.3.1   Data

The data used in the study came from the 2015/2016 Malawi Demographic and Health Survey (DHS) and the 2013–2014 Service Provision Assessments (SPA). Both surveys use

standardised questionnaires and methodology. The DHS programme provides technical support to local implementing partners of LMICs, ensuring data quality, comparability and that data is publicly available to those who need it (ICF International). The SPA and DHS surveys were conducted two years apart. Given this time difference, the focus will be mainly on comparing women's reports in the exit interviews and direct observations, since they were conducted on the same day.

The SPA survey is a census of all registered health facilities in Malawi and collected data from 977 facilities. The SPA survey included four questionnaires that are used to collect information on several aspects of quality of care. These include the following:

1. An inventory questionnaire examined the services availability and facility infrastructure;
2. The health worker questionnaire collected information on health worker demographic characteristics, responsibilities, roles and trainings. The selected sample of health workers is between 8-15.
3. The observation checklist assessed the extent to which providers adhere to clinical procedures (in antenatal care consultations, surveyors observed whether health workers conducted routine tests and prevention procedures outlined in antenatal care guidelines; the number of observations ranged between five consultations per provider and up to 15 consultations per service and were conducted in a selected facilities); and
4. Exit interview questionnaire collected information on the observed client's perception and level of service satisfaction. It also collects information on the demographic characteristics of the clients. Exit interviews are only conducted to women who accessed healthcare at the selected facilities and usually conducted immediately after the consultation.

Exit interview questionnaires were translated into Chichewa and Tumbuka. The analysis was limited to 412 facilities that reported to have antenatal care services and where randomly selected for observation consultations and client interviews. The exit interviews had a very high response rate with 99% of the clients who were observed agreeing to be interviewed. The total sample for exit interviews is 2 068 clients.

For this study, I used the Demographic and Health Survey for Malawi, conducted in 2015/2016. It is a population level household survey with national-wide data on population and health, and

has been conducted in 90 countries since 1984. Eligible households are selected using a multi-stage cluster sampling design for participation. Every woman between the ages 15-49 was eligible for participation in the women's interview and the response rate was high, over 95%. In the antenatal care section, various questions on maternal healthcare utilisation were asked to women who had a live birth in the five years prior to the survey (NSO, 2016).

The dataset for Malawi comprises of information on antenatal care for 10 192 births (with multiple pregnancies being counted as one birth) that occurred in the years 2010 and 2016. However, the main sample is limited to women who stated that they had had a live birth in the past year prior to the survey and reported to have been residing in the community for at least two years prior to giving birth. These limitations were set in order to minimise recall errors and ensure that women lived close to the clinics at the time of delivery. Furthermore, given that, literature shows that women's perception on the quality of care is dependent on the number of visits made to the provider (Kambala et al., 2015); the analysis was also limited to women who had at least three antenatal care visits[4]. The final sample for the Malawi DHS data is 807 clusters and 2 694 women.

In order to compare direct observations from the SPA with retrospective self-reports from the Malawi DHS, I linked the two datasets at the facility and cluster level using the buffer method (Mallick et al., 2017), which I describe in depth in section 4.3.3. Whereas, direct consultations and client interviews were done on the same day and at the same facility, as such, buffering was not required in linking the two datasets. Further, the comparison between the two was done at the individual level.

4.3.2 Methods

Measuring quality of care

Compared to quantity, there are various challenges when it comes to measuring quality of care, mainly because of the intricacy of quality of care (Choi, 2018; Burger et al., 2016). Quality of care has several dimensions including process, structure and outcome (Donabedian, 1980; Bruce, 1990; Jain et al., 1992).

---

[4] Nevertheless, I conducted robustness checks to make sure that the sample used for analysis was not biased as to only represent women who were more aware of the importance of ANC. I compared observation reports to women who reported to have had at least one visit. The results were the same; women in general over report the level of service quality in comparison to direct observations. Check appendix 6, Table 1.

- Structure mainly indicates the availability of certain equipment, medication, commodities and infrastructure that enables health workers to provide services (Choi, 2018).

- Process, for example, measures physical examination in antenatal care and involves technical competence and interpersonal relations (Bruce, 1990).

- Outcome measures the impact on the patient and demonstrate the result of an intervention, for example, reduced child and maternal mortality and improved child birthweight (Donabedian, 1980).

In this study, I focus on the process attributes of quality of care for the following reason: Kruk et al. (2018) argued that, unlike structure, process measures are more informative as regards the quality of care provided. This is because process measures are instantaneous and important when the client is accessing care, providing clear understanding on care provision without making any adjustments for risk (Kruk et al. 2018). This makes process attributes more appropriate in measuring gaps or inequalities in care for vulnerable subpopulations (Kruk et al., 2018). On the other hand, structure does not foretell the care that providers offer and whether that translates into improved health outcomes – more often than not, poor care occurs in the existence of sufficient tools (Leslie et al., 2017).

Selection of indicators of quality of care and coding of the questions

I identified 12 elements of antenatal care that are clinically important for the provision of antenatal care based on WHO guidelines (WHO, 2016)and also how easy they were to measure in both the DHS and SPA datasets. These elements are divided into four categories:

1. Maternal and foetal physical examination (blood pressure measured, weight measured, foetal heartbeat assessed);
2. Prevention measures (given medication for malaria prevention, given/prescribed iron/folic tablets, prescribed Albendazole and HIV-testing and counselling);
3. Screening tests (urine test for glucose and protein and blood test for syphilis and anaemia); and
4. Counselling (information given on nutrition in pregnancy, iron side effects and pregnancy complications).

Table 4.1 illustrates the domains defined in this study. Binary variables were created for each element to assess whether women received any of the mentioned antenatal care elements, using

each of the three data sources. For direct observations, researchers assessed whether or not a provider complied with antenatal care guidelines during the antenatal care consultation; for example, whether the provider conducted screening tests. I therefore coded the indicator for provision of an antenatal care service as "1" if the observer noted that the provider provided the service and '0' otherwise. In cases where the information was missing and where the provider was not observed providing the service, I coded as '0' and an assumption was made that the service was not provided (Mallick et al., 2017)[5]. For facility-level analysis and linking to the DHS clusters, I aggregated the variables to the facility level and calculated averages for each facility (Tumlinson et al., 2015). The average scores were then joined into the facility file and dichotomised as '1' if half or more than half of the women were treated, and '0' otherwise (Shmueli et.al.,2017). Despite the risk of loss of information, aggregated facility variables were dichotomised for easy comparison with the DHS variables, which were in binary format.

In the DHS, women are interviewed about the quality of antenatal care services received long after the child is born, and for the repeated routine tests and examination, women are asked whether they received the service at least once during pregnancy. I coded DHS variables as '1' if a woman reported to have received any of the antenatal care elements at least once during ANC visit, and '0' otherwise. Since the comparison between direct observations and women's DHS retrospective reports is at the facility level, I aggregated the individual responses at the cluster level to get a community-level estimate for the individual antenatal care indicators.

In the client exit interviews, women were interviewed immediately after the antenatal care visit consultation and the researcher asked whether the woman had received a particular antenatal care element in the current or the previous visit before the survey. For comparability purposes with direct observations, I only coded client exit variables as '1' if a woman reported to have either received any of the antenatal care elements in the current visit or in both the current and the previous visit[6]. The comparison between direct observation and client exit was done at individual level.

---

[5] The maximum number of values missing for the variables, was only 2 %, otherwise the majority of the variables had complete values reported.

[6] This question had four possible responses as follows; 1= yes; current visit only; 2= yes, current and previous visit; 3= previous visit only; and 4= never. I coded both 1 and 2 as 1 indicating that the service was provided and the rest 0, indicating the service was never provided.

Table 4.1: Summary of quality indicators as captured in the three assessment methods

Component	Domain	Indicator summary	Women's DHS	SPA data – Observation	SPA data – Client
Process	Physical examination	• Client's blood pressure measured • Client's weight measured • Foetal heartbeat assessed	X X X	X X X	Not captured Not captured Not captured
Process	Prevention	• Client given medication for malaria prevention • Client given/prescribed iron and/or folic tablets • Client prescribed Albendazole • Offered HIV-testing and counselling	X X X X	X X X X	X X Not captured Not captured
Process	Screening tests	• Urine test for glucose and protein • Blood test for syphilis and anaemia	X X	X X	Not captured Not captured
Process	Information given to client	• Information given to client on nutrition in pregnancy • Counselling on iron side effects • Counselling on pregnancy complications	X Not captured Not captured	X X X	X X X

Source: 2015/2016 DHS and 2013/2014 SPA

72

4.3.3   Linking facility data and individual-level data: the buffer method

The most commonly pragmatic approach in evaluating the impact of quality of care is geographically linking clinic data to households or individual data (Mallick et al., 2017). However, even though linking individuals to their closest facility is hypothetically possible, there are limits to this: household GPS coordinates are displaced in the data to maintain the anonymity of the survey participants and women sometimes bypass care from their nearest facility (Mallick et al., 2017; Montana, 2001). Skiles et al. (2013) recommend that individuals should be linked to health facility data using a measure of quality or service environment inside a broader geographic area by defining a radius for the environment relative to all localities, both urban and rural.

To minimise the displacement effects on the analysis, I followed the work of Mallick et al. (2017) and created constellations of facilities within a 10 km buffer zone of household clusters for both urban and rural to accommodate the maximum displacement of the clusters. The SPA surveys in Malawi are census surveys, which include each registered health facility across the country. This made it possible to link women to facilities at the cluster level using the created 10 km buffer distances.

Similar to the geographical linkage, it was impossible linking women to the exact facility where they accessed antenatal care. Instead, the weighted average of the individual facility elements was used to link facility data to clusters to capture the level of service provision of facilities within the 10 km buffer distance in a cluster. Clusters without any facilities offering antenatal care services within the 10KM buffer were excluded from the analysis. Of the 807 remaining clusters, 678 clusters had at least one facility offering antenatal care services within the 10 km buffer distance. Meanwhile, all 412 facilities were linked to their respective clusters.

Since direct observation and client exit interviews were conducted only in selected facilities based on the reports by healthcare providers that they offer antenatal care, I conducted descriptive analyses to compare characteristics of facilities with observation data and those without (see Appendix 5) (Mallick et al., 2017). I found that facilities without observation data were mostly lower-level facilities offering primary healthcare (non-hospitals) with antenatal care services offered for less than five days per week.

Geographically linking the observed facilities from the facility census to DHS data allowed the comparison between antenatal care quality estimates based on direct observations and retrospective self-reports based on the nationally representative DHS dataset, thus uncovering any potential discrepancies in reporting between facility estimates and retrospective women's self-reports. This is the first study to have done this.

4.3.4 Validity analysis method

Data collected through direct observations was taken as the gold standard, I was compared to the information collected through exit interviews, and retrospective self-reports in the DHS data. Two validity measures were used to conduct the analysis: sensitivity and specificity.

Sensitivity is defined as the probability that a test correctly classifies people with positive outcomes (Blanc et al. 2016). In this study's context, sensitivity is the probability of positively classifying, through exit interviews or retrospective self-reports, women who reported to have received the service; or clusters with facilities that actually offered the service, respectively. Specificity, on the other hand, is defined as the probability of correctly categorising as negative, people with negative outcomes (Blanc et al. 2016). In this study, it is the probability of correctly identifying, through exit interviews or retrospective self-reports, women who actually did not receive a particular service; or clusters with facilities that did not offer a particular antenatal care service, respectively.

Furthermore, the recall accuracy of each indicator was summarised by quantifying the area under the ROC, which plots the sensitivity (i.e. true positive rate) of each indicator against its false positive rate (1–specificity) (Blanc et al., 2016; Liu et al., 2013). I measured the uncertainty associated with the validity of each indicator by estimating 95% confidence intervals (CI), with an assumption of a binomial distribution. ROC values range between 1 (perfect classification accuracy) and 0 (zero accuracy). A ROC value of 0.5 signifies a random guess. Following the work of Blanc et al. (2016), reporting accuracy was categorised as high (ROC>0.70), moderate (0.60<ROC<0.70) and low (ROC<0.60).

Choi (2018) argued that, given that exit interviews are conducted immediately after observation in SPAs, any noteworthy inconsistencies between observation and exit interviews would therefore reflect "courtesy bias" (in cases where the estimate from the exit interview was higher) or difficulties in relaying information effectively to the client (where the estimate from

74

observation was higher), rather than recall errors. Statistical analysis was performed using Stata Version 14.

4.4   Results

4.4.1   Characteristics of healthcare facilities in the sample

A total of 412 facilities were observed, the majority of which are rural facilities (339) and the rest urban facilities (73). Among the facilities, 68% are government owned and 32% are private. A small proportion of facilities are hospitals (about 14%) whereas the majority of the facilities are health centres (81%) (Table 4.2).

Table 4.2: Characteristics of ANC facilities

Characteristics	Proportion (%)
Location of facilities	
Urban	82.28
Rural	17.72
Managing authority	
Public/government	68
Private	32
Type of facility	
Hospital	14.16
Health centre	80.83
Dispensary	2.06
Clinic	2.65
Health post	0.29

Source: Own calculations based on 2013/2014 SPA

4.4.2   Distribution of the facilities offering antenatal care services in the cluster

Figure 4.2 shows the distribution of facilities offering antenatal care services within the 10 km buffer distance in a cluster. About 16% of the clusters had no facility offering antenatal care within the 10 km buffer distance. These clusters were excluded from the analysis, leaving 678 clusters for analysis.

4.4.3   Socio-demographic characteristics of women in the sample

Table 4.3 presents the characteristics of the interviewed sample of mothers in both the client exit interviews and the DHS. Twelve per cent of the women who were interviewed during client exit interviews reported having no formal education, whereas in the DHS 11% of the women had no formal education. However, the majority of women reported to have primary education in both surveys; 64% and 66% in the client exit and DHS respondents respectively.
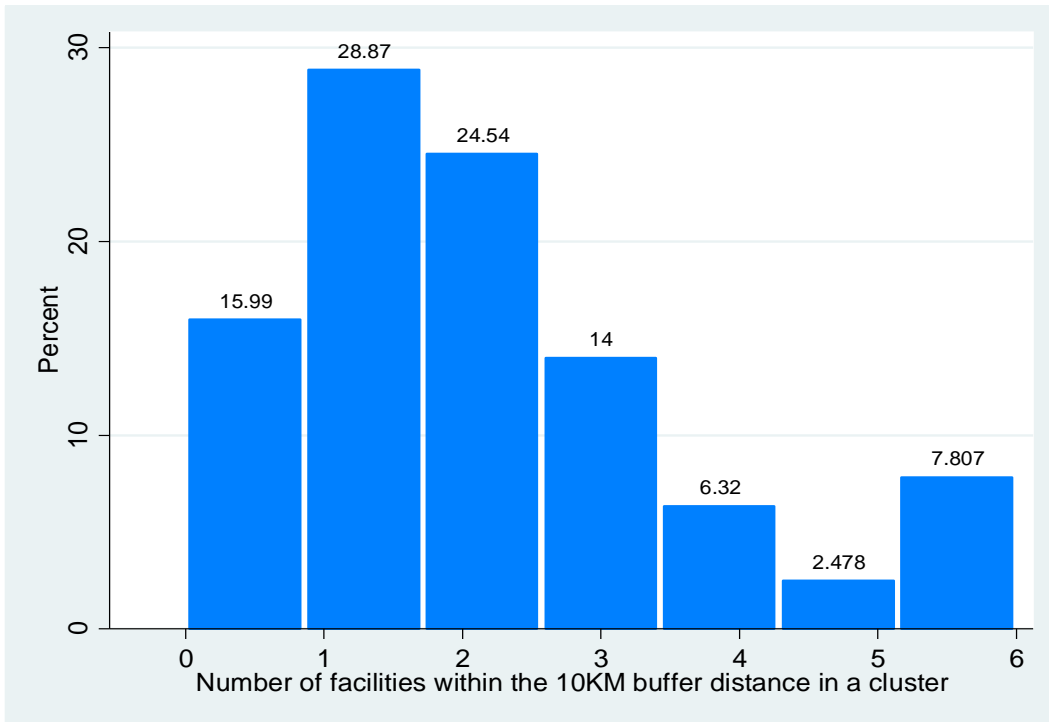
Figure 4.2: Number of facilities offering ANC services in a cluster
Source: Own calculations derived from 2015/2016 DHS and 2013/2014 SPA data

The average age at delivery was 25 and 26 for client exit and DHS respondents respectively. On average, 25% of the women in the sample are first-time mothers for both client exit interviews and the DHS. However, women in the client exit interviews were more likely to be urban residence in comparison to the DHS women. I therefore, expect women in client exit interviews to be more accurate in their reporting of quality of care given that people in urban areas are more exposed to information. Nevertheless, given a lot of similarities in the selected demographic characteristics, I view the two surveys as demographically representative of the same population.

Household and child characteristics were unavailable for the sample of women in the client exit interviews. Meanwhile, in the DHS sample, the majority of the women came from poor households and 10% of live births were low-birthweight babies. Respectively 8% and 9% of women reported having a history of Caesarean delivery and miscarriage.

Table 4.3: Selected social and demographic characteristics of the analysis sample

Variables	Client exit		DHS	
	Mean	Std. dev.	Mean	Std. dev.
Mother's education level				
None	12.46	0.33	11.32	0.32
Primary	63.53	0.48	66.78	0.47
Secondary	21.36	0.41	19.86	0.40
Tertiary	2.65	0.16	2.04	0.14
Mother's age at child's birth				
Age	24.88	6.10	26.53	6.85
Parity				
First pregnancy/child	25.31	0.43	24.80	0.43
Type of residence				
Urban	19.91	0.40	15.81	0.36
Rural	80.09	0.40	84.19	0.36
Household characteristics				
Poorest	Not available	-	22.31	0.42
Poor	Not available	-	22.87	0.42
Middle	Not available	-	18.93	0.39
Wealthy	Not available	-	18.86	0.39
Wealthiest	Not available	-	17.04	0.38
Child and pregnancy characteristics				
History of Caesarean delivery	Not available	-	7.51	0.26
History of miscarriages	Not available	-	9.13	0.29
Low birthweight	Not available	-	9.95	0.30
Number of observations	**2 068**		**2 694**	

Source: Own calculations derived from 2015/2016 DHS and 2013/2014 SPA

4.4.4 Comparison of estimates based on client exit interviews with direct observation

Direct observation and client interviews were conducted on the same day and at the same facility, therefore, any differences between them would be due to other factors other than the timing between them. Comparison between direct observation quality estimates and client exit quality estimates are shown in Figure 4.3. I identified five antenatal care elements common to both data sources. While there are no significant differences between direct observation and women's estimates on the provision or prescription of iron and folic tablets as well as medication for the prevention of malaria, for some indicators, the variation from direct observation is substantial. For example, 14% of women in the exit interviews reported that the provider(s) discussed iron side effects, while 9% of observers recorded that this took place. About 46% of providers were observed counselling women on pregnancy complications, whereas 52% of the women reported that they had received this service.

Finally, 48% of the women reported having received nutrition counselling and information on the right foods to eat during pregnancy, whereas only 37% of the observers recorded providers providing this service.
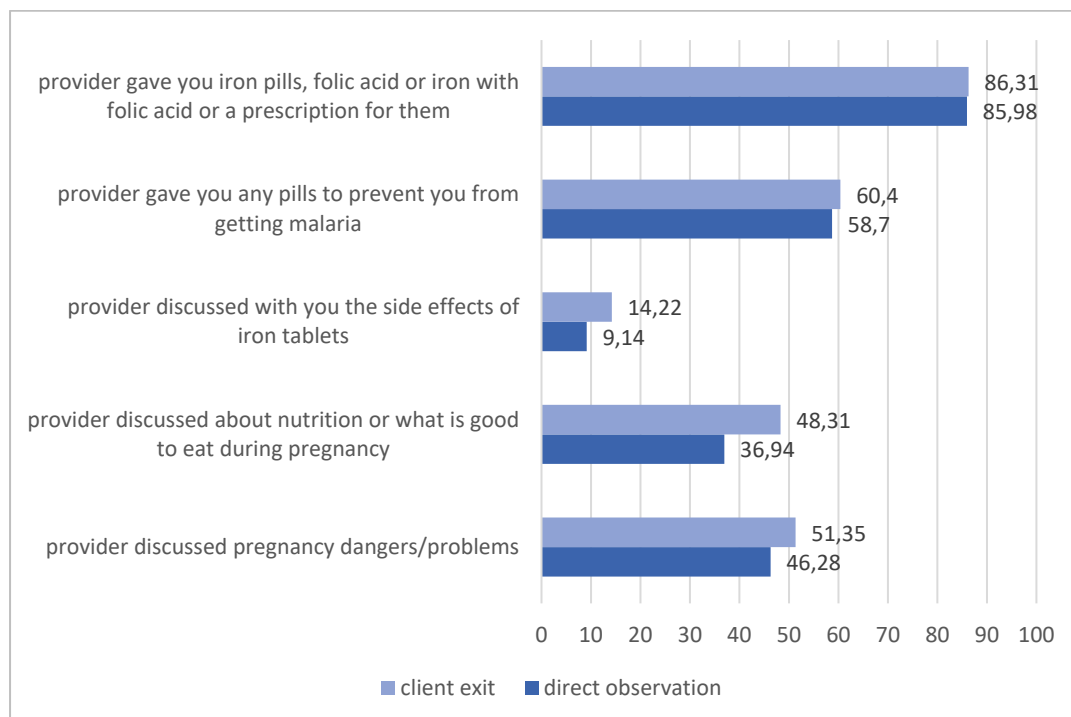
Figure 4.3: Estimates of ANC indicators by direct observation and client exit interview
Source: Own calculations derived from the 2013/2014 SPA

4.4.5 Validation of client exit estimates with direct observations

Table 4.4 presents validation results for women's reports in the client exit interviews. Sensitivity, or the knack of women to accurately report that they had received a particular antenatal care element when direct observation recorded the same, varied significantly depending on the antenatal care service. Women were more likely to report with accuracy receiving iron pills or/and folic acid (96%), medication for the prevention of malaria (87%) and counselling on nutrition and the right foods to eat during pregnancy (72%). However, women with whom the provider had discussed the side effects of iron tablets (34%) and the danger signs for complications in pregnancy (62%) were less likely recall with accuracy receiving the information from the provider.

The low sensitivity for the two indicators might be due to smaller proportions of women who were observed to have received this information, i.e. only 9% of women were observed being counselled on iron side effects and only 46% of women were observed being counselled on

pregnancy complications. It could also be due to the women's lack of knowledge and understanding of the questions. Whereas, specificity, women's ability to correctly identify not receiving a certain service, was relatively low, ranging from 59 to 83%. Women were less likely to correctly report not receiving counselling on the danger signs for pregnancy complications (57%) and nutrition information (67%). However, women were more able to correctly report not receiving counselling for iron side effects (88%). Given the low sensitivity and high specificity in the iron side effect variable, there is a possibility that women in the exit interview under-reported this service.

Looking at the accuracy in recall measure (ROC), three indicators had high reporting accuracy (ROC>70), one had moderate reporting accuracy (60>ROC<70) and another low reporting accuracy (ROC<60). Indicators with high ROC results mostly reflected objective measures of antenatal care (e.g. provision of iron and malaria prophylaxis). Indicators with low-value ROC results were mostly subjective measures and required a certain level of knowledge and understanding (counselling on pregnancy complications and iron side effects). Furthermore, I found that women with no education had relatively lower ROC values (58–78%) compared to women with primary education (58–86%) or higher than primary levels of education (60–87%). For example, women with no formal education had very low accurate recall on whether the provider had discussed the side effects of iron tablets (ROC<60%) compared to women with primary or secondary education (ROC>60%). These results indicate a positive relationship between education and the ability to recall information, meaning that it is possible that women understand and are more able to retain information given during antenatal care consultations if they are educated.

Table 4.4: Sensitivity and specificity of reporting in exit interviews compared to direct observation in SPA (%)

Survey	All clients Est.	(95% CI)	No educ Est.	(95% CI)	Primary educ Est.	(95% CI)	Secondary or higher Est.	(95% CI)
Sensitivity								
Provision of iron	95.67	(94.62 96.57)	93.25	(89.27 96.09)	95.74	(94.39 96.84)	96.87	(94.70 98.32)
Provision of prophylaxis	87.31	(85.31 89.14)	81.82	(75.31 87.22)	87.32	(84.77 89.59)	90.94	(86.82 94.11)
Advice on iron side effects	33.70	(26.91 41.02)	29.17	(12.62 51.09)	33.03	(24.32 42.69)	37.25	(24.13 51.92)
Advice on nutrition	73.82	(70.55 76.91)	72.38	(62.80 80.66)	72.61	(68.34 76.59)	77.66	(71.03 83.40)
Advice on complications	60.50	(57.32 63.61)	60.38	(50.41 69.75)	58.92	(54.85 62.91)	64.20	(58.01 70.06)
Specificity								
Provision of Iron	71.01	(65.44 76.19)	63.04	(47.55 76.79)	76.97	(70.08 82.93)	60.61	(47.81 72.42)
Provision of prophylaxis	77.87	(74.93 80.61)	68.22	(58.52 76.89)	77.97	(74.20 81.42)	80.41	(76.66 87.24)
Advice on iron side effects	87.83	(86.20 89.32)	89.96	(85.31 93.53)	89.05	(87.07 90.82)	83.29	(79.31 86.79)
Advice on nutrition	66.64	(64.01 69.20)	70.22	(62.93 76.83)	66.99	(63.68 70.18)	63.48	(57.68 69.00)
Advice on complications	56.53	(53.55 59.47)	56.50	(48.85 63.92)	56.90	(53.17 60.58)	55.36	(48.59 61.98)
ROC								
Provision of Iron	83.35	(80.70 86.01)	78.15	(70.92 85.38)	86.35	(83.20 89.51)	78.74	(72.74 84.74)
Provision of prophylaxis	82.59	(80.91 84.27)	75.02	(69.74 80.29)	82.64	(80.53 84.76)	86.68	(83.60 89.75)
Advice on iron side effects	60.76	(57.25 64.27)	59.56	(50.07 69.05)	61.04	(56.51 65.57)	60.27	(53.33 67.22)
Advice on nutrition	70.23	(68.21 72.25)	71.30	(65.84 76.76)	69.80	(67.23 72.37)	70.57	(66.50 74.64)
Advice on complications	58.51	(56.39 60.64)	58.44	(52.50 64.38)	57.91	(55.22 60.60)	59.78	(55.39 64.17)
Number of clients	**2 068**		**283**		**1 304**		**481**	

Source: Own calculations based on 2013/2014 SPA survey and 2015/2016 DH

4.4.6 Comparison between direct observation and women's retrospective reports in the DHS

The full list of indicators selected for the congruency of women's reports on the quality of care using the DHS data are presented in Table 4.5. I identified 10 indicators that were common in both the facility observations and the DHS self-reports. Using observation reports, the level of service provision in the facilities is very low, with only four indicators being provided in over 70% of the clusters. Among the lowest is the provision of anaemia and syphilis tests as well as urine testing for protein, where only 25% and 13% of the clusters had healthcare facilities providing such services respectively. These two indicators also vary significantly by location: urban clusters are more likely to have facilities that offer the two screening tests compared to rural clusters.

Table 4.5: Estimates of ANC indicators by facility observations and DHS women's reports

Services provided	Direct observation Clusters with facilities that offer the service (%)				Women's DHS reports Clusters with facilities that offer the service (%)			
	All	Urban	Rural	P-val	All	Urban	Rural	P-val
Physical examination								
Blood pressure taken	66.08	65.09	69.53	0.3091	93.95	96.03	93.36	0.2259
Client weight measured	80.09	80.79	79.89	0.8057	98.53	96.69	99.05	0.0338
Foetal heartbeat assessed	87.17	89.40	86.53	0.3522	99.26	99.43	98.68	0.3396
Screening tests								
Blood sample test for anaemia and syphilis	25.22	41.72	20.49	0.0000	98.53	98.68	98.48	0.8622
Urine tested for protein	12.83	24.50	9.49	0.0000	35.99	47.02	32.83	0.0013
Preventive interventions								
Iron and folic tablets	87.76	89.40	87.27	0.4847	98.08	98.01	98.10	0.9439
Malaria prophylaxis	76.99	80.79	75.90	0.2084	97.35	96.03	97.72	0.2536
Albendazole	51.18	45.70	52.75	0.1266	68.29	66.87	68.69	0.6752
Offered HIV test	68.73	78.81	65.84	0.0024	99.71	100	99.62	0.4491
Counselling								
Nutrition in pregnancy	54.87	61.59	52.94	0.0599	97.79	97.35	97.91	0.6796
Total sample	**678**	**151**	**527**		**678**	**151**	**527**	

Source: Own calculations derived from the 2015/2016 DHS and 2013/2014 SPA

There are significant differences between women's reports based on the services they received a year prior to the survey and reports based on facility observation. The differences are larger than when I compared facility reports to exit interviews, probably because I have added additional biases: recall and anchoring (not only lack of knowledge). For example, whereas

direct observations suggest that blood pressure measurement is conducted in 66% of the clusters, over 90% of the women reported having received this service in their cluster. This represents an overestimation of more than 20%. The same applies to the rest of the indicators: women's estimation of service provision in the clusters is significantly higher than what is suggested by the direct observations.

4.4.7   Validation of DHS self-report estimates with direct observations

Table 4.6 presents the accuracy of recall of 10 antenatal care components offered at antenatal care facilities during pregnancy at community level. In general, women tend to over-report most of the events, resulting in high sensitivity and lowered specificity and generally low ROC estimates of all items except for urine sample test. For example, based on direct observation, 66% of clusters were recorded to have facilities that provide blood pressure measurement to pregnant women, whereas women reported that 94% of the clusters had facilities that provided this service (Table 4.6), resulting in a ROC estimate of 52 (Table 4.4). Similar results were found for all the other items except for the urine test, which had a sensitivity of 47% compared to 66% specificity. This resulted in perfect reporting accuracy (ROC=1).

I also found all three validity measures comparable across women with different education backgrounds, even though communities where the majority of women had no education had comparably lower sensitivity estimates compared to communities with higher education levels. Contrary to earlier findings in the exit interviews, there were no statistical differences in the reporting accuracy of women across the three education categories. All the ROC estimates in the sample are either slightly above 0.50 or below indicating lower accuracy reporting levels on almost all recalled components of antenatal care. One possible explanation for this could be that, in the DHS retrospective data, I was not only dealing with recall bias but also with a lack of understanding/comprehension of the questions and the inability to connect the questions to what the provider did during a consultation in the distant past.

I further considered the possibility that childbirth outcomes might affect the accuracy of recall (anchoring) in retrospective data. I grouped the mothers into two categories: those who reported having a low-birthweight baby (less than 2 500 g) and those who reported having a baby weighing 2 500 g or more at birth (normal birthweight). I conducted additional analyses to test

Table 4.6: Sensitivity and specificity of reporting in direct observation in SPA compared to women's retrospective reports in DHS (%)

Survey	One-year recall			No educ			Primary			Secondary or higher		
	Est.	(95% CI)		Est.	(95% CI)		Est.	(95% CI)		Est.	(95% CI)	
Sensitivity												
Physical examination												
Blood pressure taken	95.54	(93.19	97.25)	87.88	79.78	93.58)	91.55	(88.44	94.10)	94.76	(91.03	97.26)
Client weight measured	98.34	(96.88	99.24)	99.25	(95.91	99.98)	98.75	(97.30	99.54)	98.17	(95.78	99.40)
Foetal heartbeat assessed	99.49	(80.45	88.84)	98.04	(94.38	99.59)	98.85	(97.52	99.58)	98.29	(96.05	99.44)
Screening tests												
Blood sample tested	98.25	(94.96	99.64)	96.55	(82.24	99.91)	96.62	(92.29	98.89)	98.08	(93.23	99.77)
Urine tested for protein	47.13	(36.33	58.13)	100	(81.47	100)	100	(95.14	100)	100	(93.28	100)
Preventive interventions												
Iron and folic tablets	96.37	(94.55	97.71)	87.07	(80.55	92.04)	90.40	(94.44	97.82)	95.62	(92.63	97.65)
Malaria prophylaxis	97.13	(95.30	98.38)	96.09	(91.12	98.72)	96.51	(94.40	97.99)	97.27	(94.45	98.89)
Albendazole	70.89	(65.81	75.62)	63.95	(52.88	74.03)	66.56	(61.02	71.78)	64.04	(56.53	71.09)
Offered HIV test	99.79	(98.81	99.99)	98.25	(93.81	99.79)	99.50	(98.22	99.94)	100	(98.47	100)
Counselling												
Nutrition in pregnancy	97.31	(95.11	98.70)	93.98	(86.50	98.02)	93.93	(90.68	96.31)	97.33	(93.87	99.13)

Survey	One-year recall			No educ			Primary			Secondary or higher		
	Est.	(95% CI)		Est.	(95% CI)		Est.	(95% CI)		Est.	(95% CI)	
Specificity												
Physical examination												
Blood pressure taken	9.13	(5.74	13.62)	11.84	(5.56	21.29)	13.13	(8.76	18.65)	12.86	(7.20	20.61)
Client weight measured	0.74	(0.02	4.06)	2.44	(0.06	12.86)	1.64	(0.20	5.80)	0	(0	5.52)
Foetal heartbeat assessed	8.51	(2.37	20.38)	4.55	(0.12	22.84)	1.25	(0.03	6.77)	0	(0	7.71)
Screening tests												
Blood sample tested	1.38	(0.56	2.82)	4.11	(1.52	8.73)	1.76	(0.76	3.44)	3.85	(1.77	7.18)
Urine tested for protein	65.65	(61.67	69.48)	100	(97.68	100)	100	(99.30	100)	100	(98.71	100)
Preventive interventions												
Iron and folic tablets	2.38	(0.29	8.34)	7.14	(0.88	23.50)	2.70	(0.33	9.42)	2.44	(0.06	12.86)
Malaria prophylaxis	1.92	(0.40	5.52)	12.77	(4.83	25.74)	4.20	(1.56	8.91)	7.32	(2.73	15.25)
Albendazole	34.44	(29.33	39.83)	43.82	(33.32	54.75)	32.99	(27.61	38.72)	40	(32.35	48.03)
Offered HIV test	0.47	(0.01	2.60)	0	(0	5.87)	1.01	(0.12	3.58)	1.02	(0.03	5.55)
Counselling												
Nutrition in pregnancy	1.63	(0.53	3.77)	7.61	(3.11	15.05)	3.11	(1.43	5.83)	3.97	(1.47	8.45)
ROC												
Physical examination												
Blood pressure taken	0.52	(0.50	0.54)	0.50	(0.45	0.55)	0.52	(0.49	0.55)	0.54	(0.50	0.57)
Client weight measured	0.50	(0.49	0.50)	0.51	(0.48	0.53)	0.50	(0.49	0.51)	0.49	(0.48	0.50)
Foetal heartbeat assessed	0.47	(0.42	0.51)	0.51	(0.47	0.56)	0.50	(0.49	0.51)	0.49	(0.48	0.50)
Screening tests												
Blood sample tested	0.50	(0.49	0.51)	0.50	(0.47	0.54)	0.49	(0.48	0.51)	0.51	(0.49	0.53)
Urine tested for protein	0.56	(0.51	0.62)	1	(1	1)	1	(1	1)	1	(1	1)
Preventive interventions												
Iron and folic tablets	0.50	(0.49	0.52)	0.47	(0.42	0.53)	0.50	(0.48	0.52)	0.49	(0.46	0.52)
Malaria prophylaxis	0.50	(0.48	0.51)	0.54	(0.49	0.60)	0.50	(0.49	0.52)	0.52	(0.49	0.55)
Albendazole	0.53	(0.49	0.56)	0.54	(0.47	0.61)	0.50	(0.46	0.54)	0.52	(0.47	0.57)
Offered HIV test	0.50	(0.50	0.51)	0.49	(0.48	0.50)	0.50	(0.49	0.51)	0.51	(0.50	0.52)
Counselling												
Nutrition in pregnancy	0.49	(0.48	0.51)	0.51	(0.47	0.55)	0.49	(0.47	0.50)	0.51	(0.49	0.53)
Number of clusters	**678**			**218**			**716**			**390**		

Source: Own calculations based on the 2012/2014 SPA and 2015/2016 DHS surveys

84

the hypothesis that the birthweight of a child may be negatively associated with recall accuracy (Table 4.7). I noted that, overall, women who reported having a baby with normal birthweight had significantly higher sensitivity estimates and generally lower specificity estimates, indicating a high probability of over-reporting. However, there were no substantial differences in the accuracy of reporting between the two groups of women. Generally, the ROC estimates between the two groups were similar in most cases, indicating very low accuracy levels that were equivalent to what would be generated by random guesses in both groups. The exception was urine testing, which had a perfect overlap for both specificity and sensitivity.

Table 4.7: Accuracy of recall by birthweight

Survey	Low birthweight			Normal weight		
	Estimate	(95% CI)		Estimate	(95% CI)	
Sensitivity						
Physical examination						
Blood pressure taken	89.83	(82.91	94.63)	95.53	(93.17	97.25)
Client weight measured	96.92	(92.31	99.16)	98.67	(97.28	99.46)
Foetal heartbeat assessed	96.58	(92.19	98.88)	99.13	(97.98	99.72)
Screening tests						
Blood sample tested	100	(91.78	100)	97.63	(94.05	99.35)
Urine tested for protein	100	(83.16	100)	100	(95.60	100)
Preventive interventions						
Iron and folic tablets	90.48	(84.54	94.69)	98.06	(96.55	99.03)
Malaria prophylaxis	93.94	(88.41	97.35)	96.83	(94.91	98.18)
Albendazole	65.06	(53.81	75.20)	69.71	(64.52	74.55)
Offered HIV test	97.27	(92.24	99.43)	99.05	(97.58	99.74)
Counselling						
Nutrition in pregnancy	90	(81.86	95.32)	97.27	(95.03	98.68)
Specificity						
Physical examination						
Blood pressure taken	9.80	(3.26	21.41)	11.22	(7.25	16.36)
Client weight measured	5.13	(0.63	17.32)	0.79	(0.02	4.34)
Foetal heartbeat assessed	0	(0	14.82)	1.27	(0.03	6.85)
Screening tests						
Blood sample tested	7.14	(3.32	13.13)	1.86	(0.86	3.51)
Urine tested for protein	100	(97.55	100)	100	(99.35	100)
Preventive interventions						
Iron and folic tablets	4.55	(0.12	22.84)	2.33	(0.28	8.15)
Malaria prophylaxis	13.51	(4.54	28.77)	2.72	(0.75	6.82)
Albendazole	45.35	(34.58	56.45)	33.97	(28.73	39.52)
Offered HIV test	3.39	(0.41	11.71)	0.43	(0.01	2.38)
Counselling						
Nutrition in pregnancy	10.13	(4.47	18.98)	3.50	(1.69	6.34)

85

	Low birthweight		Normal weight	
	Estimate	**(95% C.I)**	**Estimate**	**(95% C.I)**
ROC				
Physical examination				
Blood pressure taken	0.50	(0.45 0.55)	0.53	(0.51 0.56)
Client weight measured	0.51	(0.47 0.55)	0.50	(0.49 0.51)
Foetal heartbeat assessed	0.48	(0.47 0.50)	0.50	(0.49 0.51)
Screening tests				
Blood sample tested	0.54	(0.51 0.56)	0.50	(0.48 0.51)
Urine tested for protein	1	(1 1)	1	(1 1)
Preventive interventions				
Iron and folic tablets	0.48	(0.42 0.53)	0.50	(0.48 0.52)
Malaria prophylaxis	0.54	(0.48 0.60)	0.50	(0.48 0.51)
Albendazole	0.55	(0.48 0.63)	0.52	(0.48 0.55)
Offered HIV test	0.50	(0.48 0.53)	0.50	(0.49 0.50)
Counselling				
Nutrition in pregnancy	0.50	(0.45 0.55)	0.50	(0.49 0.52)
Number of observations	**201**		**776**	

4.5   Discussion of findings

Providing high-quality maternal healthcare services is essential to improve both the well-being and the survival of mothers and their newborn babies. Progress can only be monitored if the information on the care received is accurate. In Malawi, where over 95% of women have access to antenatal care with skilled personnel, the study suggests that facility-level exit interviews and retrospective women's self-reports based on household DHS on their own are not adequate to measure and monitor progress on the quality of antenatal care services provided. Using these two methods of data collection, the study found that the sensitivity, specificity and ROC area indicators of antenatal care vary considerably with direct observation. It therefore reinforces the need to compare estimates of the two data sources with direct observation of care for validation, to be able to accurately measure the quality of antenatal care.

The validation of client exit interviews against a gold standard of direct observations focusing on individual-level validity had high reporting accuracy (ROC≥70) for selected indicators about the provision of medication for the prevention of malaria, iron and/or folic acid and counselling on nutrition for a healthy pregnancy. However, client exit interviews had low reporting accuracy for counselling on the danger signs for complications in pregnancy and the side effects of iron. Since the estimates from the clients were relatively higher than from direct observation, the discrepancy in the reports between the data sources could be due to the following two factors:

1. **Courtesy bias:** Given that client exit interviews are conducted at the facility, the overestimation in the level of service provision by the women may reflect courtesy bias. Courtesy bias happens when a client is reluctant to share negative experiences due to the proximity of service providers. This is well documented in the literature (Avis et al., 1997; Simmons & Elias, 1994).

2. **Lack of understanding:** The second reason could be a lack of understanding and knowledge concerning the questions asked by the researcher. It appears that, where the researcher did not simplify the question or could not further simplify the question, and where sophisticated terminology was used, such as 'side effects' or 'complications', the accuracy was remarkably lower. For example, questions related to counselling on the side effects of iron and on pregnancy-related complications had lower reporting accuracy compared to straightforward questions that asked whether the provider had prescribed medication for malaria prevention or iron/folic acid tablets.

Nevertheless, the results are not surprising. Similar findings have been reported by Gupta et al. (2011), who found low validity of exit interviews for counselling in the management of common childhood diseases in Afghanistan, and Choi (2018), who found low validity of exit interviews for counselling in family planning in selected LMICs. Furthermore, validation of client exit interviews also shows significant differences in reporting accuracy by education level. I found that women with no education had relatively lower ROC values (58–78%) compared to women with primary or higher levels of education (60–87%). These results indicate a positive relationship between education and the ability to recall information, which means it is possible that women in client exit interviews understand better, what happens in antenatal care consultations if they are better educated. Gupta et al. (2011) also found that the woman's ability to accurately recall of her past varies with the level of her education.

The study also adds new evidence to the validity of retrospective household survey data, where women are asked about pregnancy related information and antenatal care utilisation long after they have already given birth. Even though the DHS interviews women who had a live birth five years prior to the survey, I focused on women who gave birth a year prior to the survey to minimise errors in recall. The results show that women's retrospective self-reports had very high sensitivity and lower specificity estimates, indicating significant overestimation of the level of service quality by the women. This resulted in very low accuracy in recall, as indicated

by low ROC values ranging between 0.47 and 0.54 across all 10 antenatal care components. The ROC estimates represent significant recall errors and random guesses in the responses of the women. The two-year gap between the DHS and the SPA facility census and differences in the timing of the two surveys, as previously highlighted, made it difficult to synthesise conclusions. For example, the DHS assesses the quality of care over the entire period of pregnancy, whereas facility direct observations assess the quality of care at one point in time, in this case, on the day of the survey.

Nevertheless, besides potential recall errors, the combination of high sensitivity and low specificity in the women's self-reports may also reflect social desirability amongst women, based on their expectation of receiving appropriate care (Kruk et al., 2018). That is, if a woman sees that it is socially desirable to receive a certain service, it increases the likelihood of them reporting to having received the service, even if they did not essentially receive it (Kruk et al., 2018). Other researchers have also argued that women from vulnerable groups, such as the rural poor, tend to have lower expectations of service provision. Olatunde et al. (2018) argued that the low expectations of the objectively poor, often times lead to higher quality ratings. This could be the case in this study, as the majority of the facilities and women are from rural communities.

Other studies have also discussed additional reasons why women may perceive the quality of care differently to the actual observed facility care. In their study of errors in recalling childhood socio-economic status in South Africa, Von Fintel and Posel (2016) explained that another problem that has potential to affect the consistency and dependability of recall data is connected to anchoring. Haas (2007) elucidated that anchoring occurs when the respondents' current circumstance or status, influence their retrospective reports. Contrary to the client exit interviews, in the DHS women report on care after the child is born, therefore there is a possibility that they could display a cognitive bias by overstating the care they received in response to the positive life experience that they had. If they had reported this before the birth, when there was more uncertainty regarding the pregnancy, there would be no anchoring. Confirming this argument is a study by Blanc et al. (2016) which found that the validity of women's reports of care they received during birth in Kenya was influenced by whether the woman had a Caesarean section. In this study, however, I did not find any association between

pregnancy outcomes (low birthweight) and the accuracy of women's reports on antenatal care indicators.

The usage of direct observation of care as the gold standard strengthens this study, given that clinical records are of poor quality in low-income settings. However, direct observations could be susceptible to error from changing behaviour by the health worker (the Hawthorne effect) (I did not measure the extent of the Hawthorne effect in this study), even when quality control mechanisms are there (Masatu et al., 2006). Nevertheless, relatively lower prevalence rates as measured by direct observations in comparison to the other two methods mean that probable overestimation owing to the Hawthorne effect may not offset the strengths of the method for policy purposes to effectively measure the level of services provision and improve the quality of antenatal care services. Additional strengths include the geographical linking of facilities to household surveys, which allowed conducting the validity of women's reports on antenatal care quality by using not only client exit interviews but also retrospective self-reports based on household survey data.

However, the reader should take into account several limitations when interpreting the findings of this study.

1. The SPA includes observation data from a sample of facilities among the censured facilities. Besides, clients who were observed and interviewed included both those who had just initiated ANC for the first time and returning follow up clients. This means the provider may not have conducted all the tests and examinations for follow up clients, as they would for new clients (Choi, 2018).

2. The timing of the interviews differs for direct observations and DHS data. In the direct observations, it is not clear for how many antenatal care visits the woman was observed. Although gestation weeks for the pregnancy were captured, I could not disaggregate the data by the gestation weeks due to sample limitation. On the other hand, the DHS interviews women about the antenatal care services long after the pregnancy and only looks at whether a service had been provided at least once during the entire pregnancy period. Given the highlighted major differences in the timing between the two data sources, the discrepancies between them are too large to be convincingly explained by assumptions necessary for the merge or mapping done in this study.

89

3. The DHS surveys only focused on women who had a live birth prior to the survey. This means that women who had a miscarriage or a still birth, or whose child died a few months before the survey, and had received poor-quality antenatal care, were not included in the survey. As such, the DHS data may overestimate the level of service provision.

4.6   Conclusion

Applying accurate methods of assessing the performance of health workers is critical to measuring antenatal care service quality in Malawi and other LMICs, particularly where the burden of maternal and child mortality is so large, and more vulnerable women have access to basic maternal health services. It is, therefore, vital to understand the extend of any prejudices that each particular quality assessment method has and discrepancies in comparison to other methods in order to avoid erroneous conclusions and to select a combination of methods most suitable to the issue being assessed (Blanc et al. 2016).

In this study, I found that client exit interviews are limited by the mothers' lack of knowledge about complicated components of care, for example counselling on the side effects of iron and the danger signs for complications in pregnancy. This is shown in the different patterns by education and the highest ROC scores for the most educated women. This means that, in cases where mothers are educated, recall of complicated elements of antenatal care would be reported more accurately and exit interviews would be more reliable. I therefore, recommend the use of client exit interviews to evaluate the quality of healthcare; however, where indicators require a certain level of knowledge and understanding, the use of direct observations to validate client exit interviews is highly recommended.

In LMICs, facility surveys are often lacking, and data on process measures of quality is usually scarce (Choi, 2018). Population-based surveys might fill the gap in data due to their wide availability; however, there are limitations in using retrospective self-reported data. The likelihood of recall errors is very high, since participants are asked to make a recall of their past (i.e. up to five years for maternal health services). In the case of this study, evidence of recall errors was seen even among women who had given birth only a year prior to the survey. Therefore, when monitoring maternal health services using standard household surveys,

policymakers should consider comparing and using multiple data sources, including direct facility observation data and where possible clinical records.

Overall, this work has the potential to improve the understanding of the measurement of quality of antenatal care and thus also, hopefully, placing these measures prominently in policy dialogues and targets in LMICs.

CHAPTER FIVE

Summary and conclusion

5.1   Summary of findings

For the past decade, the government of Malawi has made many health sector reforms. The aim has been – and still is – "to improve the quality of life of the entire population by reducing the risk of ill health and the occurrence of premature deaths" (MoH, 2017). In that respect, Malawi has over the past decade both implemented and introduced a number of preventive maternal healthcare interventions in an effort to improve maternal and child health outcomes. However, the impact and the effectiveness of these interventions is yet to be seen, accordingly Malawi continues to register the worst maternal and child health outcomes in comparison to other countries in the sub Saharan region. For example, high antenatal coverage rates, remarkably well over 90%, continue to coincide with comparatively high maternal mortality rates. Given the country's failure to attain the maternal health-related Millennium Development Goal (MDG) of reducing the maternal mortality rate by 75% in 2015, the relevance of maternal and child health to the UN's new Sustainable Development Goals as well the overarching Malawi Health Sector Strategic Plan II (2017–2022), there is a need to illuminate the factors that influence the supply and demand of preventive maternal healthcare interventions.

In this regard, the thesis analysed three main research objectives:

- The first objective was to evaluate the performance of the World Health Organization's (WHO) 2001 Focused Antenatal Care (FANC) policy in Malawi.
- The second objective was to estimate the optimal number of antenatal care visits that are effective in improving birth outcomes in low-resourced settings.
- The final objective was to assess the validity of women's self-reports on the quality of antenatal care services received during pregnancy against the gold standard of direct client observations at facilities.

I addressed the objectives empirically, using various datasets and approaches, in three separate individual chapters.

The initial analysis chapter (chapter two) was a retrospective study, focused on examining the impact of the 2001 FANC policy on the early access to, underutilisation of and quality of antenatal care services. Based on data pooled from three comparable nationally representative Malawi Demographic and Health Survey (DHS) datasets (2000, 2004 and 2010) and using the interrupted time series analysis, I found that FANC is associated with earlier access to care. However, it has also been associated with unintended increases in underutilisation. I saw no change in the quality of antenatal care services. Finally, there were significant disparities in FANC performance in rural and urban areas, with the urban areas performing better.

Effective implementation of FANC required significant investment in equipment and human resources (Lungu et al., 2011); however, the government of Malawi did not make the necessary changes in this regard. The underinvestment by the government in the required structures for successful FANC implementation potentially explains why the FANC policy did not translate into improved quality of care. Dupas (2011) argued that the provision of a high quality of care could have a substantial influence on health-seeking behaviour. As such, policies that aim to strengthen the capacity of the health system, for example provision of more training in FANC guidelines and engaging community health workers to educate rural women on the importance of seeking antenatal care early and utilising FANC services, could help in the effective implementation of and increased demand for FANC services.

In 2015, results from a systematic review reported that the reduced FANC model was linked with an increased risk of perinatal mortality in comparison to a model with a minimum number of eight visits. This prompted a revision of the FANC model by the WHO in 2016, when new antenatal care guidelines were issued and the minimum number of visits was doubled from four to eight. Given that Malawi and other low- and medium-income countries (LMICs) alike have failed to effectively implement the less ambitious four-visit model, it is therefore questionable whether they would be successful in implementing the new guidelines. Villar et al. (2001) argued that most antenatal care policies in LMICs are adopted without thorough scientific evaluation and evidence-based programming on the optimal number of visits. I addressed this topic as part of the second research objective.

In the second objective (chapter three), I estimated the optimal number of visits that are effective in reducing low birthweight in Malawi. Using the nationally representative 2015/2016

Malawi DHS data, I applied instrumental variable models together with highly flexible spline specifications and Wald tests to estimate breaks in the relationship between the number of antenatal care visits and the probability of low birthweight. In so doing, I established thresholds at which antenatal care stops mattering for the reduction of low birthweight. I also explored the relationship between the number of visits and the quality of care and their independent influences on low birthweight. I found a structural break at three antenatal care visits, beyond which the marginal benefit of an additional antenatal care visit is zero. Unlike the previous studies, which used dose-response analysis or qualitative methods to show changes in birth outcomes across different models of antenatal care, the study used quasi-experimental methods. I have therefore been able to show a causal relationship between the number of visits and low birthweight.

Furthermore, even though the bivariate analysis shows a linear relationship between the number of visits and the quality of care from three up to eight visits, I did not find any evidence of a direct pathway for quality of care on low birthweight. Based on these findings, a reduced model of antenatal care may perform just as well for Malawi and additional visits may then be recommended for women who have complicated pregnancies to reduce the risk of perinatal mortality cases. Evidence shows that perinatal mortality cases are higher in women with risky pregnancies than in women with normal and uncomplicated pregnancies (Evers et al., 2010). However, due to data limitations, I was unable to determine the number of visits based on the risk profile of the women in this study. Nevertheless, I controlled for variables such as a history of miscarriages, Caesarean deliveries and the age of the mother at delivery in all the regressions as risk proxies. Lastly, other factors such as the mother's education and household economic status had an independent influence on low birthweight. Therefore, policies that seek to empower women educationally can also be effective in improving birth outcomes.

In addressing the first two research objectives, I took advantage of the publicly available nationally representative household Demographic and Health Surveys, which are based on women's self-reports on service utilisation and quality of care. However, the reliability of this data depends on a number of factors. Among others, these include: the client's ability to recall with accuracy (memory); how much attention the client paid to the provider's actions; the client's access to information and knowledge of the content of the component of care; an understanding of the questions being asked and the ability to link them to what the provider

94

was doing; the willingness to participate, and how comfortable the client was talking to an interviewer (Franco et al., 2002). Furthermore, a client's self-reports may be affected by their recent condition and socio-demographic characteristics (anchoring) (Von Fintel & Posel, 2015). The highlighted limitations may lead to an upward or downward bias in the quality of care measurement or services utilisation, which has implications for monitoring the quality of care. In chapters two and three, an effort was made to minimise some of the highlighted weaknesses by limiting the samples to the most recent birth prior to the surveys. However, as described, there are other potential biases, which require further exploration and are, therefore, addressed with the third objective (chapter four).

Despite the documented weaknesses of self-reported data in LMICs, very few studies have validated women's reports on the quality of care against facility estimates, particularly due to a lack of reliable facility data. In the fourth chapter, I took advantage of the availability of both facility census data and household data to compare women's estimates of the quality of care against facility estimates. Addressing the third research objective, I used direct observations of ANC consultations as the reference standard and compared them to exit interviews with pregnant women after antenatal care consultation and retrospective self-reports on the provision of antenatal care services as captured in the DHS. I specifically tested the sensitivity, specificity and receiver-operating curves (ROCs) of antenatal care quality indicators as captured in the three methods of data collection.

The main result of interest was the comparison between client exit interviews and direct observations as they were conducted on the same day. I found that client exit interviews are limited by the mothers' lack of knowledge about complicated components of care, for example counselling on the side effects of iron and the danger signs for complications in pregnancy. Women in low-resourced settings lack basic information on the specific interventions they must expect when they do access healthcare. If they do have the information, they may not be able to process it and understand what is happening, mostly due to lower education levels (Dupas, 2011). In the study, this is shown by different reporting patterns based on the education levels of the women. Women with higher levels of education had the highest ROC scores, indicating better understanding, and high reporting accuracy compared to women with lower levels of education. This means that, in cases where mothers are educated, the recall of

complicated elements of antenatal care would be reported accurately and exit interviews would be reliable.

In addition, I found that retrospective self-reports are even further from observations than exit interviews, and are highly inflated. I found evidence of recall errors in women's reports even among women who had given birth a year prior to the survey. However, given the differences in the gestational point in time when the questions were administered in the DHS and direct observations, the discrepancies are too large to be convincingly explained by assumptions made in this study. The difference in the timing of the two data is more likely to have lead to upward bias and hence inflated estimates in the women's DHS reports. Nevertheless, the findings suggest that, when monitoring maternal health services using standard household surveys, policymakers should consider comparing and using multiple data sources, including direct facility observations and clinical records if available to get accurate estimates of quality of care.

5.2   Study limitations

While much has been uncovered about the effectiveness of antenatal care models and the methods of measuring maternal healthcare in Malawi, using available data, the reader should take note of several limitations. Similar to Chama-Chiliba's (2013) experience in her study of maternal health care utilisation in Zambia, one of the major concerns in this study was issues related to the use of secondary data, such as the quality level of data and the matching of the available data, finding the right match between the data used and the problem being investigated. This particularly is related to the comparison between DHS retrospective self-reports and facility direct observations. The timing of these two surveys are different. Direct observations are conducted at one specific time during the pregnancy, whereas the self-retrospective reports are conducted after the child is born, and assess quality of care over the entire pregnancy. I therefore had to aggregate both the facility and the household data before linking the two datasets to make the comparison.

Additionally, in the case of facility data, while a majority of the facilities covered in the census offer antenatal care services, direct observations and client exit interviews were not conducted in all facilities that offered antenatal care services, creating selection bias if the selection of the facilities was non-random. I conducted a comparison between observed facilities and non-observed facilities and found that most of the non-observed facilities were of lower levels

(mostly primary) and only provided antenatal care less than five times in a week. Another issue of concern comes from the displacement of the DHS coordinates to protect the identity of the DHS respondents, as it breeds noise in the linking of households to the facilities and to the rainfall instrument. Furthermore, women who died due to childbirth-related complications or whose babies died prior to the DHSs were not included in this analysis. This, therefore, indirectly assumes that more maternal deaths and poor birth outcomes are associated with inadequate use of antenatal care and poor antenatal care quality. There is thus a possibility that this analysis overestimated the utilisation of antenatal care services, quality of care and the impact of FANC.

Finally, the findings in paper 3 that women in the DHS are more likely to overestimate the level of service quality have implications on the findings of the first two papers. There is a possibility that the quality of care measures in the first two papers, were measured with error leading to a downward bias. However, using an aggregated measure of quality potentially controlled for such measurement error in our analysis.

5.3   Policy implications and conclusion

Notwithstanding the limitations in the analysis, the thesis has contributed to a better understanding of the barriers to maternal health interventions in Malawi and the obstacles that may impede the achievement of the new Sustainable Development Goals. As previously stated, the country's principal objective is to "ensure improved quality of life for its entire population, including maternal care, in the hope of improving health outcomes". However, these efforts have been hampered by a lack of empirical evidence. This thesis contributes to addressing this need for more evidence to guide policy reforms.

The results have shown that, when it comes to policy implementation, one size does not always fit all. I found that the successful implementation of antenatal care policies hinges upon the policy environment and the capacity strength of the health system, which, for countries like Malawi, is often a challenge. I saw that, despite FANC being a cost-effective model having required a minimum of only four visits, the model had no significant impact on the quality of care and had the unintended consequence of increasing underutilisation of antenatal care. The results support the need for policies that will strengthen the capacity of health systems, that will make them more efficient, and that will ensure that maternal healthcare policies are

implemented properly. This requires a rigorous effort based on empirical evidence to scale up interventions so that more people can benefit from these impacts.

Related to the need for policy implementation based on empirical evidence, the results show that only the first three antenatal care visits are associated with a reduction in low birthweight in Malawi. These results have two implications for preventive antenatal care policies in low-resourced settings and the WHO's 2016 revision of the minimum number of visits from four to eight.

1. Given the limited resources in Malawi, limiting the number of visits to three or maintaining a reduced-visit model, may be the best option to make use of the constrained resources wisely, and maximise quality of care (subject to the constraints).

2. If Malawi is to adopt the 2016 guidelines, it may require that the capacity of the health system be expanded because only then can the new guidelines be successfully implemented. However, given that Malawi has failed to successfully implement the four-visit model due to a lack of proper infrastructure and capacity, this latter option may be far-fetched. The eight visits may work where there is a lot of infrastructure and capacity to provide sophisticated care at different junctures during the pregnancy. However, where those resources are lacking, there is no point in doubling the number of visits.

5.4   Suggestions for future research

There are some areas that this thesis did not cover, due to limitations in time and data, but which are imperative in the Malawian context. For instance, related to chapter three, there is a need to further explore the average number of antenatal care visits that are effective for improving other maternal and child health outcomes, for example neonatal deaths, maternal deaths and the likelihood of pre-eclampsia. Data on these aspects was not available for this study. I also suggest a multilevel country study, which will provide a lot of variation in the number of antenatal care visits, to estimate the average number of visits effective for improved birth outcomes in LMICs. Another aspect worthy of further analysis is using primary data and qualitative methods to analyse the specific components of antenatal care that are the most beneficial in improving both maternal and child health outcomes.

Related to chapter four, it would be interesting to conduct a follow-up study in a year or two to follow up with women who were interviewed in the exit interviews to assess whether indeed a longer recall period has an impact on reporting accuracy. In the study, I used DHS data to validate women's retrospective reports, which I based on the assumption that women went to facilities within their 10 km buffer distance. Therefore, any significant differences between DHS retrospective reports and direct observations may have been due to factors other than recall errors.

Regarding the supply aspect (quality of care) of maternal healthcare, the overall findings from this thesis suggest that effective maternal health policy interventions are only possible if strategies are formulated to strengthen the existing structures and reduce system inefficiencies. Regarding the demand aspect (utilisation of care), there is a need for strategies that increase access to information on the benefits of antenatal care and that empower women through education. As Michael Grossman (1972) explained, education is a factor that affects the demand for healthcare: a more educated person may be more able to maintain good health and make more effective use of healthcare and preventive services.

Most importantly, the findings suggest that public policy can play an important role when it comes to preventive healthcare. Demand-side policy tools such as increased access to basic information on the importance of accessing antenatal care services can only be successful if the supply side is adequate and effective (Dupas, 2011).

BIBLIOGRAPHY

Abdal Qader MA, Badilla I, Mohd Amin R, Ghazi HF (2012) Influence of antenatal care on birth weight: A cross sectional study in Baghdad City Iraq. *BMC Public Health.*

Acharya, D., Khanal, V., Singh, J. K., Adhikari, M., & Gautam, S. (2015). Impact of mass media on the utilization of antenatal care services among women of rural community in Nepal. *BMC Res Notes, 8*, 4–9.

Adekanle, D.A., & Isawumi, A. I. (2008). Late antenatal care booking and its predictors among pregnant women in south-western Nigeria. *Online Journal of Health and Allied Sciences, 7*, 1–6.

Ahmed, F. U., & Das, A. M. (1992). Beneficial effects. Three ANC visits might be the divergent point in lowering low birthweight babies. *Integration, 33*, 50–53.

Alan, J. G. (1980). Parametric methods in the analysis of survival data. *Microelectron Reliab, 20*, 477–481.

Alkema, L., Chou, D., Hogan, D., et al. (2016). Global, regional, and national levels and trends in maternal mortality between 1990 and 2015, with scenario-based projections to 2030: A systematic analysis by the UN Maternal Mortality Estimation Inter-Agency Group. *The Lancet, 387*, 462–474.

Al-Samarrai, S., & Zaman, H. (2006). Abolishing school fees in Malawi: The impact on education access and equity. MPRA Paper No. 130. Retrieved on 30 June 2018 from http://mpra.ub.uni-muenchen.de/130/

Angrist, J. D. (2001). Estimation of limited dependent variable models with dummy endogenous regressors: Simple strategies for empirical practice. *Journal of Business & Economic Statistics, 19(1)*, 2–16.

Arrow, K. J. (1963). Uncertainty and the welfare economics of medical care. *The American Economic Review, 53(5)*, 941–973.

Atuoye, N. K., Dixon, J., & Luginaah, I. (2015). Can she make it? Transportation barriers to accessing maternal and child health services in rural Ghana. *BMC Health Services Research, 15*, 333.

Aung, T., Dominic, M., Karen, S., Thin, M. K., & Willi, M. (2012). Validation of a new method for testing provider clinical quality in rural settings in low and middle-income countries:

The observed simulated patient. *PLOS ONE, 7(1)*, e30196. Retrieved from http://dx.plos.org/10.1371/journal.pone.0030196

Avis, M., Bond, M., & Arthur, A. (1997). Questioning patient satisfaction: An empirical investigation in two outpatient clinics. *Soc Sci Med, 44(1)*, 85–92.

Awiti, J.O. (2014). A multilevel analysis of prenatal care and birth weight in Kenya. *Health Economics Review, 4*, 33. Retrieved from doi: 10.1186/s13561-014-0033-3

Ayinde, O. O., Oladeji, D. B., Abdulmalik, J., Jordan, K., Kola, L., & Gureje, O. (2018). Quality of perinatal depression care in primary care setting in Nigeria. *BMC Health Services Research, 18*, 879.

Bbaale, E. (2011). Factors influencing utilization of antenatal care content in Uganda. *Am Med J, 4(9)*, 516–526. Retrieved from http//dx.doi.org/10.4066/AMJ.2011.849.

Bellows, B., Behl, R., Abuya, T., Muriuki, A. M., Bajracharya, A., & Choi, Y. (2016). Benchmarking to assess quality of family planning services: Construction and use of indices for family planning readiness in Kenya with data from 2010 and 2014. In A. Sprockett, S. H., Leisher, K., Longfield., & D. Montagu (Eds.), *Quality Measurement in Family Planning: Past, Present, Future.* Papers from the Bellagio Meeting on Family Planning Quality. Oakland, California, USA: Metrics for Measurement.

Benova, L., Tunçalp, Ö., Moran, C. A., & Campbell, O. (2018). Not just a number: Examining coverage and content of antenatal care in low-income and middle-income countries. *BMJ Glob Health*.

Berglund, A. C., & Lindmark, G. (1998). Health services effects of a reduced routine programme of antenatal care: An area based study. *Eur J Obstet Gynecol Reprod Biol, 77*, 193–199.

Bhalotra, S. (2010). Fatal fluctuations? Cyclicality in infant mortality in India. *Journal of Development Economics, 93*, 7–19.

Bhattacharya, A. A., Allen, E., Umar, N., Usman, A. U., Felix, H., Audu, A., Joanna, R. M., Schellenberg, J. R. M., & Marchant, T. (2019). Monitoring childbirth care in primary health facilities: A validity study in Gombe State, northeastern Nigeria. *Journal of Global Health, 9(2)*. Retrieved from doi: 10.7189/jogh.09.020411

Birungi, H., & Onyango-Ouma, W. (2006). Acceptability and sustainability of the WHO Focused Antenatal Care package in Kenya. *Frontiers in Reproductive Health.* Washington D C: Population Council.

Blanc, A. K., Warren, C., McCarthy, K. J., Kimani, J., Ndwiga, C., RamaRao, S. (2016). Assessing the validity of indicators of the quality of maternal and newborn healthcare in Kenya. *Journal of Global Health, 6(1)*. Retrieved from doi: 10.7189/jogh.06.010405.

Boller, C., Wyss, K., Mtasiwa, D., & Tanner, M. (2003). Quality and comparison of antenatal care in public and private providers in the United Republic of Tanzania. *Bull World Health Organ, 81(2)*, 116–122.

Bonell, C. P., Hargreaves, J., Cousens, S., et al. (2009). Alternatives to randomisation in the evaluation of public health interventions: Design challenges and solutions. *J Epidemiol Community Health, 65*, 582–587.

Brown, L., Tyane, M., Bertrand, J., Lauro, D., Abou-ouakil, M., & deMaria, L. (1995). Quality of care in family planning services in Morocco. *Stud Fam Plann, 26(3)*, 154–168.

Bruce, J. (1990). Fundamental elements of the quality of care: A simple framework. *Studies in Family Planning, 21(2)*, 61–91. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2191476

Burger, R., Ranchod, S., Rossouw, L., & Smith, A. (2016). Strengthening the measurement of quality of care. *SAHR*.

Burgert, C. R., & Prosnitz, D. (2014). Linking DHS Household and SPA Facility Surveys: Data considerations and geospatial methods. DHS Spatial Analysis Reports No. 10. Rockville, Maryland, USA: ICF International. Retrieved from http://dhsprogram.com/pubs/pdf/SAR10/SAR10.pdf

Buscha, F., & Conte, A. (2009). A bivariate ordered probit estimator with mixed effects. Jena Economic Research Papers. ISSN 1864-7057.

Carter, E., Tuuli, G. M., & Cahill, G. A. (2016). Number of prenatal visits and pregnancy outcomes in low-risk women. *J Perinatol, 36(3)*, 178–181.

Casey, R., Rieckhoff, M., Beebe, S. A., & Pinto-Martin, J. (1992). Obstetric and perinatal events: The accuracy of maternal report. *Clin Pediatr (Phila), 31*, 200–204.

Chama-Chiliba, C. M. (2013). An economic analysis of maternal healthcare in Zambia. Retrieved in May 2016 from https://books.google.co.za

Chama-Chiliba, C. M., & Koch, S. F. (2015). Utilization of Focused Antenatal Care in Zambia: examining individual- and community-level factors using a multilevel analysis. *Health Pol Planning, 30*, 78–87. Retrieved from doi: 10.1093/heapol/czt099

Chege, J. (2005). *Feasibility of introducing a comprehensive integrated package of antenatal care services in rural public clinics in South Africa*. Washington DC: Population Council, Frontiers in Reproduction Health.

Chen, L. F., Carriker, C., Staheli, R., Isaacs, P., Elliott, B., Miller, B. A., et al. (2013). Observing and improving hand hygiene compliance: Implementation and refinement of an electronic-assisted direct-observer hand hygiene audit program. *Infect Control Hosp Epidemiol, 34*, 207–210.

Choi, Y. (2018). Estimates of Side effects counselling in family planning using three data sources: Implications for monitoring and survey design. *Studies in Family Planning, 49(1)*.

Crossley, T. F., & Kennedy, S. (2001). The reliability of self-assessed health status. *Journal of Health Economics, 21*, 643–658.

Da Fonseca, C. R., Strufaldi, M.W., de Carvalho, L. R., Puccini, R. F. (2014). Adequacy of antenatal care and its relationship with low birth weight in Botucatu, Sao Paulo, Brazil: A case-control study. *BMC Pregnancy and Childbirth*. Retrieved from https://doi.org/10.1186/1471-2393-14-255

Dale, H., & Schuler, S. R. (1993). The simulated client method: Evaluating client-provider interactions in family planning clinics. *Studies in Family Planning, 24(3)*, 187–193. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8351699

Das, J., Hammer, J., & Leonard, K. (2008). The quality of medical advice in low and middle-income countries. *Journal of Economic Perspectives, 22(2)*, 93–114.

Dembélé, M., & Zwart, S. J. (2016). Evaluation and comparison of satellite-based rainfall products in Burkina Faso, West Africa. *International Journal of Remote Sensing, 37(17)*, 3995–4014. Retrieved from doi: 10.1080/01431161.1207258

Do, M. A., Micah, L., Brondi, H., Campbell, T., Marchant, T. E., & Munos, M. (2016). Linking Household and Facility Data for better coverage measures in reproductive, maternal, newborn, and child healthcare: Systematic review. *Journal of Global Health, 6(2)*, 020501.

Donabedian, A. (1980). *Explorations in Quality Assessment and Monitoring: The Definition of Quality and Approaches to its Assessment*. Ann Arbor, MI: Health Administration Press.

Dowswell, T., Carroli, G., Duley, L., Gates, S., Gülmezoglu, A. M., Khan-Neelofur, D., & Piaggio, G. (2015). Alternative versus standard packages of antenatal care for low-risk pregnancy. *Cochrane Database of Systematic Reviews, 7*, Art. No.: CD000934. Retrieved from doi: 10.1002/14651858.CD000934.pub3

Dupas, P. (2011). Health behavior in developing countries. UCLA & NBER working paper prepared for the *Annual Review of Economics, 3*.

Edward, B. (2011). Factors influencing the utilisation of antenatal care content in Uganda. *Australas Med J, 4(9)*, 516.

Elkadry, E., Kenton, K., White, P., Creech, S., & Brubaker, L. (2003). Do mothers remember key events during labour? *Am J Obstet Gynecol, 189*, 195–200. Retrieved from Medline: 12861162 doi: 10.1067/mob.2003.371

Evers, A. C., Brouwers, H. A., Hukkelhoven, C. W., Nikkels, P. G., Boon, J., et al. (2010). Perinatal mortality and severe morbidity in low and high risk term pregnancies in the Netherlands: Prospective cohort study. *BMJ, 2*, 341.

Fagbamigbe, A. F., & Idemudia, E. S. (2015). Assessment of quality of antenatal care services in Nigeria: Evidence from a population-based survey. *Reproductive Health, 12*, 88. Retrieved from doi 10.1186/s12978-015-0081-0

Feikin, D. R., Audi, A., Olack, B., Bigogo, G. M., Polyak, C., Burke, H., Williamson, J., & Breiman, R. F. (2010). Evaluation of the optimal recall period for disease symptoms in home-based morbidity surveillance in rural and urban Kenya. *International Journal of Epidemiology, 39(2)*, 450–458. Retrieved from http://ije.oxfordjournals.org/content/39/2/450

Filmer, D., Hammer, J. S., Pritchett, L. S. (2000). Weak links in the chain: A diagnosis of health policy in poor countries. *The World Bank Research Observer, 15(2)*, 199–224.

Filmer, D., & Pritchett, L. (1999). The impact of public spending on health: Does money matter? *Soc Sci Med, 49(10)*, 1309–1323.

Franco, L. M., Franco, C., Kumwenda, N., et al. (2002). Methods for assessing quality of provider performance in developing countries. *Int J Qual Healthcare, 14 (1)*, 17–24.

Frankenberg, E. (1995). The effects of access to healthcare on infant mortality in Indonesia. *Health Transition Review, 5*, 143–163.

Furuta, M., & Salway, S. (2006). Women's position within the household as a determinant of maternal healthcare use in Nepal. *Int Fam Plan Perspect, 32(1)*, 17–27.

Gage, A. J. (2007). Barriers to utilisation of maternal healthcare in rural Mali. *Social Science and Medicine, 65(8)*, 1666–1682.

Gage, A. J., & Zomahoun, D. L. (2017). Influence of the service delivery environment on family planning outcomes in Nigeria. *African Population Studies, 31(1)*.

Gajate-Garrido, G. (2013). The impact of adequate prenatal care on urban birth outcomes: An analysis in a developing country context. *Economic Development and Cultural Change, 62(1)*, 95–130.

Gamedze-Mshayisa, D. I., Kuo, S., Liu, C., & Lu, Y. (2018). Factors associated with women's perception of and satisfaction with quality of intrapartum care practices in Swaziland. *Midwifery, 57*, 32–38.

Gilchrist, V. J., Strange, K. C., Flocke, S. A., et al. (2004). A comparison of the National Ambulatory Medical Care Survey (NAMCS) Measurement Approach with direct observation of outpatient visits. *Med Care, 42*, 276–280.

Gissler, M., & Hemminki, E. (1994). Amount of antenatal care and infant outcome. *Eur J Obstet Gynecol Reprod Biol, 56(1)*, 9–14.

GIZ Health. (2011). Sustainable structures for the health sector in Malawi: Cost study of service level agreement (SLA) covered services. GIZ Health, Lilongwe.

Gortmaker, L. S. (1979). The effects of prenatal care upon the health of the newborn. *American Journal of Public Health, 69(7)*.

GoM. Government of Malawi. (2014). 2014 Millennium Development Goal report for Malawi. Government of Malawi. Lilongwe, Malawi.

GoM. Government of Malawi. (2017). Health Sector Strategic Plan II 2017–2022: Towards universal health coverage. Lilongwe, Malawi.

Government Spending Watch. (2015). 2014 planned spending by the Government of Malawi. GSW, London. Retrieved from http://www.governmentspendingwatch.org/spending-data

Global Fund. (2019). Retrieved from https://www.theglobalfund.org/en

Graham, W. J., & Varghese, B. (2012). Quality, quality, quality: Gaps in the continuum of care. *The Lancet, 379(9811)*, e5-6.

Greenacre, M., & Blasius, J. (2006). Correspondence analysis and related methods in practice. In M. Greenacre, J. Blasius (Eds.), *Multiple Correspondence Analysis and Related Methods*. Chapman and Hall, London.

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

Grossman, M., & Joyce, T. (1990). Unobservable, pregnancy resolutions, and birth weight production functions in New York City. *Journal of Political Economy, 98(5)*, 983–1007.

Gudayu, T. W., Woldeyohannes, S. M., & Abdo, A. A. (2014). Timing and factors associated with first antenatal care booking among pregnant mothers in Gondar Town, North West Ethiopia. *BMC Pregnancy and Childbirth, 14(1)*, 33.

Gupta, R., & Talukdar, B. (2017). Frequency and timing of antenatal care visits and its impact on neonatal mortality in EAG states of India. *J Neonatal Biol, 6*, 263. Retrieved from doi: 10.4172/2167-0897.1000263. ISSN: 2167-0897

Gupta, S., Yamada, G., Mpembeni, R., Frumence, G., Callaghan-Koru, J. A., Stevenson, R., et al. (2014). Factors associated with four or more antenatal care visits and its decline among pregnant women in Tanzania between 1999 and 2010. *PLOS ONE, 9(7)*.

Gupta, V., Gu, K., Chen, Z., Lu, W., Ou Shu, X., & Zheng, Y. (2011). Concordance of self-reported and medical chart information on cancer diagnosis and treatment. *BMC Medical Research Methodology, 11*, 72.

Haas, S. A. (2007). The long-term effects of poor childhood health: An assessment and application of retrospective reports. *Demography, 44(1)*, 113–135.

Haessler, S. (2014). The Hawthorne effect in measurements of hand hygiene compliance: A definite problem, but also an opportunity. *Br Med J, 23*, 965–967.

Hameed, W., Ishaque, M., Gul, X., & Siddiqui, J. (2018). Does courtesy bias affect how clients report on objective and subjective measures of family planning quality? A comparison between facility- and home-based interviews. *Journal of Contraception, 9*, 33–43.

Hermida, J., Nicholas, D. D., & Blumenfeld, S. N. (1999). Comparative validity of three methods for assessment of the quality of primary healthcare. *Int J Qual Healthcare, 11*, 429–433.

Hitimana, R., Lindholm, L., Krantz, G., Nzayambaho, M., & Pulkki-Brannstrom, M. (2018). Cost of antenatal care for the health sector and for households in Rwanda. *BMC Health Services Research, 18*, 262.

Hodgins, S. (2013). Achieving better maternal and newborn outcomes: Coherent strategy and pragmatic, tailored implementation. *Glob Health Sci Pract, 1*, 146–153. Retrieved from Medline: 25276527 doi: 10.9745/GHSP-D-13-00030

Hutchinson, L. P., Do, M., & Agha, S. (2011). Measuring client satisfaction and the quality of family planning services: A comparative analysis of public and private health facilities in Tanzania, Kenya and Ghana. *BMC Health Services Research, 11*, 203.

Ibrahim, J., Yorifuji, T., Tsuda, T., Kashima, S., & Doi, H. (2012). Frequency of antenatal care visits and neonatal mortality in Indonesia. *Journal of Tropical Paediatrics, 58(3)*.

Jackson, E. F., Siddiqui, A., Gutierrez, H., Kanté, A.M., Austin, J., & Phillips, J.F. (2015). Estimation of indices of health service readiness with a principal component analysis of the Tanzania Service Provision Assessment Survey. *BMC Health Services Research, 15(1)*, 536. Retrieved from http://ovidsp.ovid.com

Jain, A., Bruce, J., & Mensch, B. (1992). Setting standards of quality in family planning programs. *Studies in Family Planning, 23*, 392–395. Retrieved from http://www.jstor.org/stable/1966897

Jayachandran, V., Chapotera, G., & Stones, W. (2016). Quality of facility-based family planning services for adolescents in Malawi: Findings from a National Census of Health Facilities. *Malawi Medical Journal, 28(2)*, 48–52.

Jensen, R. (2000). Agricultural volatility and investments in children. *American Economic Review, 90(2)*, 399–404.

Jewell, D., Sharp, D., Sanders, J., & Peters, T. J. (2000). A randomised controlled trial of flexibility in routine antenatal care. *BJOG, 107*, 1241–1247.

Jewell, T. R & Triunfo, P. (2006). The impact of prenatal care on birthweight: The case of Uruguay. *Health Econ, 15*, 1245–1250. Retrieved from doi: 10.1002/hec.1121

Johnson, K. B., Jacob, A., & Brown, M. E. (2013). Forest cover associated with improved child health and nutrition: Evidence from the Malawi Demographic and Health Survey and satellite data. *Glob Health Sci Pract, 1(2)*.

Joshi, C., Torvaldsen, S., Hodgson, R., & Hayen, A. (2014). Factors associated with the use and quality of antenatal care in Nepal: A population-based study using the Demographic and Health Survey data. *BMC Pregnancy Childbirth, 14*, 94.

Kabakian-Khasholian, T., Campbell, O., Shediac-Rizkallah, M., & Ghorayeb, F. (2000). Women's experiences of maternity care: Satisfaction or passivity? *Soc Sci Med, 51(1)*, 103–113.

Kambala, C., Lohmann, J., Mazalale, J., Brenner, S., De Allegri1, M., Muula, A. S., & Sarker M. (2015). How do Malawian women rate the quality of maternal and newborn care? Experiences and perceptions of women in the central and southern regions. *BMC Pregnancy and Childbirth, 15*, 169. Retrieved from doi: 10.1186/s12884-010560-x

Kawungezi, P. C., Akiibua, D., Aleni, C., Chitayi, M., Niwaha, A., Kazibwe, A., et al. (2015). Attendance and utilization of antenatal care (ANC) services: Multi-center Study in upcountry areas of Uganda. *Open J Prev Med, 5(3)*, 132–142.

Khanal, V., Adhikari, M., & Gavidia, T. (2014). Factors associated with the utilisation of postnatal care services among the mothers of Nepal: Analysis of Nepal Demographic and Health Survey 2011. *BMC Women's Health, 14(19)*.

Khatun, S., & Rahman, M. (2008). Quality of antenatal care and its dose-response relationship with birthweight in a maternal and child health training institute in Bangladesh. *J Biosoc Sci, 40(3)*, 321–337.

Kigenyi, O., Tefera, G. B., Nabiwemba, E., & Orach, C. G. (2013). Quality of intrapartum care at Mulago national referral hospital, Uganda: Clients' perspective. *BioMed Central Pregnancy and Childbirth, 13*, 1–8. Retrieved from http://dx.doi.org/10.1186/1471-2393-13-162

Koenraads, M., Phuka, J., Maleta, K., Theobald, S., & Gladstone, M. (2017). Understanding the challenges to caring for low birthweight babies in rural southern Malawi: A qualitative study exploring caregiver and health worker perceptions and experiences. *BMJ Glob Health, 2(3)*, e000301. Retrieved from doi: 10.1136/bmjgh-2017-000301.eCollection 2017

Kramer, M. S. (1987). Determinants of low birthweight: Methodological assessment and meta-analysis. *Bull World Health Organ, 65(5)*, 663–737.

Kruk, M., Gage, A., Arsenault, C., Jordan, K., Leslie, H. H., et al. (2018). High-quality health systems in the Sustainable Development Goals era: Time for a revolution. *The Lancet Glob Health*.

Kuhnt, J., & Vollmer, S. (2017). Antenatal care services and its implications for vital and health outcomes of children: Evidence from 193 surveys in 69 low-income and middle-income countries. *BMJ Open, 7*: e017122.

Kulkarni, M., & Nimbalkar, M. (2008). Influence of socio-demographic factors on the use of antenatal care. *Indian J Soc Prev Med, 39(3&4)*, 98–102.

Kyei, N. N. A, Chansa, C., & Gabrysch, S. (2012) Quality of antenatal care in Zambia: A national assessment. *BMC Pregnancy and Childbirth, 12*, 151. Retrieved from http://www.biomedcentral.com/1471-2393/12/151

Lawn, J.E., Blencowe, H., Waiswa, P., Amouzou, A., Mathers, C., Hogan, D., et al. (2016). Stillbirths: Rates, risk factors, and acceleration towards 2030. *The Lancet 6(387)*, 587–603.

Larroque, B., Bertrais, S., Czernichow, P., & Leger, J. (2001). School difficulties in 20-year olds who were born small for gestational age at term in a regional cohort study. *Pediatrics, 108*, 111–115.

Leonard, K., & Masatu, M. C. (2006). Outpatient process quality evaluation and the Hawthorne effect. *Soc Sci Med, 63*, 2330–2340.

Leslie, H. H., Sun, Z., & Kruk, M. E. (2017). Association between infrastructure and observed quality of care in 4 healthcare services: A cross-sectional study of 4,300 facilities in 8 countries. *PLOS Med, 14*: e1002464.

Levin, A., Dmytraczenko, T., Mceuen, M., Ssengooba, F., Mangani, R., & Van Dyck, G. V. (2003). Costs of maternal health care services in three Anglophone African countries. *International Journal of Health Planning and Management, 18*, 3–22.

Lindelow, M. W. (2003). *Health Facility Surveys: An Introduction*. Washington, DC. The World Bank.

Lino, Y., Sillabutra, J., & Chompikul, J. (2011). Factors related to the perception of pregnant women regarding antenatal care in Nakonpathom province, Thailand. *Jour of Pub Hea and Dev, 9*, 105–116.

Liu, L., Li, M., Yang, L., et al. (2013). Measuring coverage in MNCH: A validation study linking population survey derived coverage to maternal, newborn, and child healthcare records in rural China. *PLOS One, 8*, e60762.

Lungu, F., Malata, A., Chirwa, E., & Mbendera, I. (2011). Quality assessment of focused antenatal care services in Malawi. *Afr J Midwifery Women's Health, 5(4)*, 169–175.

Machira, K., & Palamuleni, M. (2018). Women's perspectives on quality of maternal healthcare services in Malawi. *International Journal of Women's Health, 10*, 25–34. Retrieved from http://dx.doi.org/10.2147/IJWH.S144426

Madden, J. M., Quick, J. D., Ross-Degnan, D., et al. (1997). Undercover care seekers: Simulated clients in the study of health provider behaviour in developing countries. *Soc Sci Med, 45*, 1465–1482.

Majoko, F., Munjanja, S. P., Nystrom, L., Mason, E., & Lindmark, G. (2007). Randomised controlled trial of two antenatal care models in rural Zimbabwe. *BJOG, 114(7)*, 802–811.

Makaula, P., Bloch, P., Banda, T. H., Mbera, B. G., Mangani, C., Sousa, A., Nkhono, E., Jemu, S., & Muula, S. A. (2012). Primary health care in rural Malawi – a qualitative assessment exploring the relevance of the community-directed interventions approach. *BMC Health Services Research, 12*, 328.

Malata, A. (2016). Safe motherhood initiative – the case of Malawi. Retrieved from https://www.wilsoncenter.org/sites/default/files/address_malata_safe_motherhood_ini tiative-_the_case_of_malawi.pdf

Mallick, L., Wang, W., & Temsah, G. (2017). A comparison of summary measures of quality of service and quality of care for family planning in Haiti, Malawi, and Tanzania. DHS Methodological Report No. 20. USAID.

Malloy, M. H., Kao, T., & Lee, Y. J. (1992). Analysing the effect of prenatal care on pregnancy outcomes: A conditional approach. *American Journal of Public Health, 82*, 448–450.

Mamba, K. C., Muula, A. S., & Stones, W. (2017). Facility-imposed barriers to early utilization of Focused Antenatal Care services in Mangochi District, Malawi – a mixed methods assessment. *BMC Pregnancy and Childbirth, 17*, 44. Retrieved from doi: 10.1186/s12884-017-1631-y

Maternal Health Task Force. (2014). Funding facility established to accelerate progress to end preventable maternal and child death. Retrieved in January 2019 from www.mhtf.org

Mayo, E. (1933). The social problems of an industrial civilization. Retrieved on 13 September 2019 from https://archive.org/stream/socialproblemsof00mayo/ socialproblemsof00mayo_djvu.txt

McCormick, M. C. (1985). The contribution of low birth weight to infant mortality and childhood morbidity. *New England Journal of Medicine, 312*, 82–90.

McDuffie, R. S., Beck, R., Bischoff, K., Cross, J., & Orleans, M. (1996). Effect of frequency of prenatal care visits on perinatal outcome among low-risk women. *JAMA, 275*, 847–885.

Mchenga, M., Burger, R., & Von Fintel, D. (2019). Examining the impact of WHO's Focused Antenatal Care policy on early access, underutilisation and quality of antenatal care services in Malawi: A retrospective study. *BMC Health Services Research, 19*, 295.

McKenzie, D. J. (2006). Disentangling age, cohort and time effects in the additive model. *Oxford Bulletin of Economics and Statistics, 68(4)*, 0305-9049.

Mensch, B., Arends-Kuenning, M., & Jain, A. (1996). the impact of the quality of family planning services on contraceptive use in Peru. *Studies in Family Planning, 27(2)*, 59–75.

MoH. Ministry of Health. (2003). FANC Guidelines. Lilongwe, Malawi.

MoH. Ministry of Health. (2007). Roadmap for accelerating the reduction of maternal and neonatal mortality and morbidity in Malawi. Lilongwe, Malawi.

MoH. Ministry of Health. (2010). Assessment Report for Health facilities in Malawi in the Implementation of FANC Services. Lilongwe, Malawi.

MoH. Ministry of Health. (2014a). Malawi National Heath Accounts with Subaccounts for HIV/AIDS, Malaria, Reproductive Health, and Child Health for Financial Years 2009/10, 2010/11, and 2011/12. Ministry of Health / Department of Planning and Policy Development: Lilongwe.

MoH. Ministry of Health. (2014b). Malawi Service Provision Assessment 2013–2014 report. Lilongwe, Malawi. Malawi Ministry of Health and Population.

Montana, L., Bessinger, R., & Curtis, S. (2001). Linking Health Facility and Population Level Data in Kenya Using Geographic Information Systems. Calverton, Maryland, USA: ORC Macro.

Munjanja, S. P., Lindmark, G., & Nystrom, L. (1996). Randomised controlled trial of a reduced-visits programme of antenatal care in Harare, Zimbabwe. *The Lancet, 348*, 364–369.

Muula, A. S. (2006). Shortage of health workers in the Malawian public health services system: How do parliamentarians perceive the problem? *African Journal of Health Sciences, 13*, 124–130.

Mwabu, G. (2009). The production of child health in Kenya: A structural model of birth weight. *J Afr Economies, 18*, 212–260.

Naomi, S., Yuri, I., & Akiko, I. (2009). Gender differences in stomach cancer survival in Osaka, Japan; Analyses Using Relative Survival Model. *Japan J Clin Oncol, 10*, 690–694.

NSO. National Statistical Office. (2005). Third Integrated Household Survey 2004–2005: Household Socio-Economic Report. Zomba, Malawi: NSO.

NSO. National Statistics Office. (2016). Malawi Demographic and Health Survey report, 2015/2016. Zomba, Malawi NSO: ICF Macro.

NSO. National Statistics Office. (2018). 2018 Population and Housing Census report. Zomba, NSO.

Ndawo-Brumblay, S. K., Mbaruku, G., & Kruk, M. E. (2012). Parity and institutional delivery in rural Tanzania: A multilevel analysis and policy implications. *Health Policy and Planning*. Retrieved from doi: 10.1093/healpol/czs104

Nwaru, B. I., Wu, Z., & Hemminki, E. (2012). Determinants of the use of prenatal care in rural China: The role of care content. *Maternal and Child Health Journal, 16(1)*, 235–241. Retrieved from https://doi.org/10.1007/s10995-010- 0734-0 PMID: 21184157

Nyarko, P., Birungi, H., Arma-Klemesu, M., Arhinful, D., Deganus, S., Odoi-Agyarko, H., & Brew, G. (2006). Acceptability and feasibility of introducing the WHO Focused Antenatal Care package in Ghana. *Frontiers in Reproductive Health.* Washington DC: Population Council.

Oladapo, O. T., Iyaniwura, C. A., & Sule-Odu, A. O. (2008). Quality of antenatal services at the primary care level in southwest Nigeria. *Afr J Reprod Health, 12*, 71–92.

Olatunde, O. A., Bibilola, D. O., Abdulmalik, J., Jordan, K., Kola, L., & Gureje, O. (2018). Quality of perinatal depression care in primary care setting in Nigeria. *BMC Health Services Research, 18*, 879. Retrieved from doi: 10.1186/s12913-018-3716-3

Olsen, R. J. (1980). A least squares correction for selectivity bias. *Econometrica, 48*, 1815–1820.

Pell, C., Meñaca, A., Were, F., Afrah, N.A., Chatio, S., et al. (2013). Factors affecting antenatal care attendance: Results from qualitative studies in Ghana, Kenya and Malawi. *PLOS ONE, 8(1)*: e53747. Retrieved from doi: 10.1371/journal.pone.0053747

Petrou, S., Kupek, E., Vause, S., & Maresh, M. (2001). Clinical, provider and socio-demographic determinants of the number of antenatal visits in England and Wales. *Soc Sci Med, 52*, 1123–1134.

Prual, A., Toure, A., Huguet, D., & Laurent, Y. (2000). The quality of risk factor screening during antenatal consultations in Niger. *Health Policy Planning, 15(1)*, 11–16.

Quigley, M. A., Hockley, C., & Davidson, L. L. (2007). Agreement between hospital records and maternal recall of mode of delivery: Evidence from 12,391 deliveries in the UK Millennium Cohort Study. *BJOG, 114*, 195–200. Retrieved from Medline: 17166217 doi: 10.1111/j.1471-0528.2006.01203

Rani, M., Bonu, S., & Harvey, S. (2008). Differentials in the quality of antenatal care in India. *International Journal of Quality Healthcare, 20*, 62–71.

Risnes, K. R., Vatten, L. J., Baker, J. L., Jameson, K., Sovio, U., Kajantie, E., et al. (2011). Birthweight and mortality in adulthood: A systematic review and meta-analysis. *Int J Epidemiol, 40*, 647–661. Retrieved from doi: 10.1093/ije/dyq267

Roberts, J., Marshak, H. H., Sealy, D. A., Taylor, L. M., Mataya, L., & Gleason, P. (2016). The role of cultural beliefs in accessing antenatal care in Malawi: A qualitative study. *Public Health Nursing*, 0737-1209. Wiley Periodicals, Inc. Retrieved from doi: 10.1111/phn.12242

Rosenzweig, M. R., & Schultz, T. P. (1983). Estimating a household production function: Heterogeneity, the demand for health inputs, and their effects on birth weight. *The Journal of Political Economy, 91*, 723–746.

Rudman, A., El-Khouri, B., & Waldenström, U. (2007). Women's satisfaction with intrapartum care: A pattern approach. *Journal of Advanced Nursing, 59*, 474–487.

Sagna, M. L., & Sunil, T. S. (2012). Effects of individual and neighbourhood factors on maternal care in Cambodia. *Health and Place, 18*, 415–423.

Say, L., & Raine, R. (2007). A systematic review of inequalities in the use of maternal healthcare in developing countries: Examining the scale of the problem and the importance of context. *Bull World Health Organ, 85*, 812–819.

Schultz, T. P. (1984). Studying the impact of household economic and community variables on child mortality. *Population and Development Review Supplement: Child Survival: Strategies for Research 10*, 215–235.

Sen, D. K., Yusof, K., & Rajaswari, K. (1991). Are three antenatal care visits enough for low risk mothers? *Singapore J Obstet Gynaecol, 22*, 83–86.

Shmueli, G., Bruce, C. P., Yahay, I., Patel, R. N., & Lichtendahl, C. K. (2017). Data Mining for Business Analytics: Concepts, Techniques, and Applications in R. *Wiley,* 1st edition.

Sikorski, J., Wilson, J., Clement, S., Das, S., & Smeeton, N. (1996). A randomised controlled trial comparing two schedules of antenatal care visit: The antenatal care project. *BMJ, 312*, 546–553.

Simkhanda, B., Teijlingen, E. R. V., Porter, M., & Simkhanda, P. (2008). Factors affecting the utilization of antenatal care in developing countries: A systematic review of the literature. *J Advanced Nurs, 61(3)*, 244–260.

Simmons, R., & Elias, C. (1994). The study of client-provider interactions: A review of methodological issues. *Stud Fam Plann, 25*, 1–17.

Skiles, M. P., Burgert, C. R., Curtis, S. L., & Spencer, J. (2013). Geographically Linking Population and Facility Surveys: Methodological Considerations. *Population Health Metrics, 11(1)*, 14.

Sofaer, S., & Firminger, K. (2006). Patient perceptions of the quality of health services. *Annul Rev Public Health, 26*, 513–559.

Soumerai, S. B. (2015). How do you know which healthcare effectiveness research you can trust? A guide to study design for the perplexed. *Prev Chronic Dis, 12*, E101.

Stanton, C. K., Rawlins, B., Drake, M., dos Anjos, M., Cantor, D, et al. (2013). Measuring Coverage in MNCH: Testing the validity of women's self-report of key maternal and newborn health interventions during the peripartum period in Mozambique. *PLOS ONE, 8*, 5.

Stewart, M. K., & Festin, M. (1995). Validation study of women's reporting and recall of major obstetric complications treated at the Philippine General Hospital. *International Journal of Gynaecology & Obstetrics, 48*, 853–866. Retrieved on 29 December 2015 from http://www.ijgo.org/article/002072929502320C/fulltext

Tafere, T. E., Afework, M. F., & Yalew, A. W. (2018). Provider's adherence to essential contents of antenatal care services increases birth weight in Bahir Dar City Administration, North West Ethiopia: A prospective follow up study. *Reproductive Health, 15*, 163. Retrieved from https://doi.org/10.1186/s12978-018-0610-8

Testa, J., Ouédraogo, C., Prual, A., De Bernis, L., & Koné, B. (2002). Determinants of risk factors associated with severe maternal morbidity: Application during antenatal consultations. *J Gynécol Obstét Biol Reprod (Paris), 31*, 44–50.

Titaley, C. R., Dibley, M. J., Roberts, C. L. (2010). Factors associated with underutilization of antenatal care services in Indonesia: Results of Indonesia Demographic and Health Survey 2002/2003 and 2007. *BMC Public Health, 10*, 485.

Troude, P., L'Hélias, L. F., Raison-Boulley, A. M., Castel, C., Bouyer, J., & De la Rouchebrochard, E. (2008). Perinatal factors reported by mothers: Do they agree with medical records? *Eur J Epidemiol, 23*, 557–564. Retrieved from Medline: 18560979 doi: 10.1007/s10654-008-9268-9

Tumlinson, K., Pence, B. W., Curtis, S. L., Marshall, S. W., & Speizer, I. S. (2015). Quality of care and contraceptive use in urban Kenya. *Int Perspectt Sex Reprod Health, 41(2)*, 69–79.

UNICEF. (2018). 2018/19 Health Budget Brief Towards Universal Health Coverage (UHC): Achieving SDG 3 in Malawi. Lilongwe, UNICEF.

UNICEF Data. (2017). Monitoring the situation of children and women: Only half of the women worldwide receive the recommended amount of care during pregnancy. Retrieved from http://www.data.unicef.org/maternal-health/antenatal-care

UN. United Nations. (2016). The 2015 Sustainable Development Goals. Retrieved on 12 January 2017 from http://www.un.org/sustainabledevelopment/health.shtml

UN IGME. United Nations Inter-agency Group for Child Mortality Estimation Report. (2017). Levels & Trends in Child Mortality. Estimates developed by the UN Inter-agency Group for Child Mortality Estimation, United Nations Children's Fund, New York.

Valera, C., Young, S., Mkandawire, N., Groen S. R., Banza, L., & Viste, A. (2019). Transportation barriers to access health care for surgical conditions in Malawi: A cross sectional nationwide household survey. *BMC Public Health, 19(264)*.

Van den Broek, N., Ntonya, C., Kayira, E., et al. (2005). Preterm birth in rural Malawi: High incidence in ultrasound-dated population. *Hum Reprod, 20*, 3235–3237.

Victora, C. G., Habicht, J. P., & Bryce, J. (2004). Evidence-based public health: Moving beyond randomized trials. *Am J Public Health, 94*, 400–405.

Villar, J., Ba'aqeel, H., Piaggio, G., Lumbiganon, P., Miguel, B. J., Farnot, U., et al. (2001). WHO antenatal care randomized trial for the evaluation of a new model of routine antenatal care. *The Lancet, 357(9268)*, 1551–1564.

Von Fintel, D., & Posel, D. (2016). Errors in recalling childhood socio-economic status: The role of anchoring and household formation in South Africa. *Soc Indic Res, 126*, 119–140. Retrieved from doi: 10.1007/s11205-015-0896-7

Vyas, S., & Kumaranayake, L. (2006). Constructing socio-economic status indices: How to use principle components analysis. *Health Policy and Planning, 21(6)*, 459–468.

Wang, W. R., Winter, L., Mallick, L., Florey, C., et al. (2015). The relationship between the health service environment and service utilization: Linking population data to health facilities data in Haiti and Malawi. DHS Analytical Studies No. 51. Rockville, Maryland, USA: ICF International. Retrieved from http://dhsprogram.com/pubs/pdf/AS51/AS51.pdf

Weeks, A., & Temmerman, M. (2016). New WHO antenatal care model – quality worth paying for? *The Lancet, 388*, 2574–2575.

Wehby, L. G., Jeffrey, C., Murray, B., Eduardo, E., Castilla, C., Jorge, S., et al. (2009). Prenatal care demand and its effects on birth outcomes by birth defect status in Argentina. *Economics and Human Biology,* 84–95.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. The MIT Press, Cambridge.

World Bank. (2015). 2015 World Development Indicators. World Bank. Retrieved on 1 July 2018 from http://data.worldbank.org/country/Malawi

World Food Programme. (2018). Logistics Capacity Assessment tool. Retrieved from http://dlca.logcluster.org/display/public/DLCA/2.3+Malawi+Road+Network;jsessionid=BA

WHO. World Health Organization. (2002a). WHO antenatal care randomized trial: Manual for the implementation of the new model. WHO document WHO/RHR/01.30. WHO, Geneva.

WHO. World Health Organisation. (2002b). Global action for skilled attendants for pregnant women. Geneva: Department of Reproductive Health and Research, World Health Organization. Retrieved from http://whqlibdoc.who.int/hq/2002/WHO_RHR_02.17.pdf

WHO. World Health Organisation. (2005). The World Health Report: Make Every Mother and Child Count. Geneva, WHO.

WHO. World Health Organization. (2008a). Global Health Observatory: Healthcare worker density. Geneva, WHO.

WHO. World Health Organization. (2008b). Framework and standards for country health information systems. Retrieved from http://www.who.int/healthmetrics/ documents/hmn_framework200803.pdf

WHO. World Health Organization. (2009). WHO recommended interventions for improving maternal and newborn health. Geneva: WHO Department of Making Pregnancy Safer.

WHO. World Health Organization. (2011). The Abuja Declaration: Ten Years On. Geneva, WHO.

WHO. World Health Organization. (2012). Comprehensive implementation plan on maternal, infant and young child nutrition: Sixty-fifth world health assembly Geneva.

WHO. World Health Organization. (2014a). WHO Global Health Workforce Statistics. Retrieved on 1 July 2018 from http://www.who.int/hrh/statistics/hwfstats/en

WHO. World Health Organization. (2014b). The implementation of option B+ for Prevention of Mother-To-Child Transmission of HIV: The Malawian experience. Brazzaville: WHO Regional Office for Africa.

WHO. World Health Organisation. (2015a). Trends in maternal mortality: 1990 to 2015: Estimates by WHO, UNICEF, UNFPA, World Bank Group and the United Nations Population Division. World Health Organization. Retrieved on 23 March 2017 from http://www.who.int/reproductivehealth/publications/monitoring/maternalmortality 2015/en

WHO. World Health Organization. (2015b). Health Expenditure Indicators. WHO, Geneva. Retrieved from http://apps.who.int/nha/database/Select/Indicators/en

WHO. World Health Organization. (2016). World Health Organization recommendations on antenatal care for a positive pregnancy experience. Geneva: WHO.

WHO. World Health Organization. (2017). WHO global database on maternal health indicators, 2017 update. Geneva: World Health Organization. Retrieved from http://www.who.int/gho/maternal_ health/en

Yaya, S., Bishwajit, G., Ekholuenetale, M., Shah, V., et al. (2017). Timing and adequate attendance of antenatal care visits among women in Ethiopia.

PLOS ONE, 12(9), e0184934. Retrieved on 30 June 2018 from https://doi.org/10.1371/journal.pone.0184934

Zamawe, O. F. C., Banda, M., & Dube, N. A. (2016). The impact of a community driven mass media campaign on the utilisation of maternal health care services in rural Malawi. *BMC Pregnancy and Childbirth, 16*, 21.

Zhou, H., Wang, A., Huang, X., Guo, S., Yang, Y., Martin, K., Tian, X., et al. (2019). Quality antenatal care protects against low birth weight in 42 poor counties of Western China. *PLOS ONE, 14(1)*: e0210393. Retrieved from https://doi.org/10.1371/journal.pone.0210393

Zhu, H. P., Xia, X., Adnan, A., Liu, S. F., & Du, K. Y. (2011). Application of Weibull model for survival of patients with gastric cancer. *BMC Gastroenterology, 11*, 1. Retrieved from https://doi.org/10.1186/1471-230X-11-1

APPENDICES

Appendix 1: Sample characteristic for reported and unreported birthweight

Table 1: Sample characteristic for reported and unreported birthweight

Covariates	Reported	Not reported	P-value
Antenatal care			
Average number of antenatal care visits	3.72	3.30	0.0000
Child characteristics			
Boy	50.00	50.00	0.1601
Birth order			
1st born	25.74	23.89	0.0450
2–4th born	50.82	46.27	0.0000
Higher older	23.44	29.84	0.0000
Multiple births			
Twin	35.17	57.93	0.0000
Maternal characteristics			
Age at birth	26.03	26.02	0.9825
Married	83.65	82.89	0.3274
Education level			
No education	10.84	21.70	0.0000
Primary	65.63	69.86	0.0000
Secondary	21.63	8.14	0.0000
Tertiary	1.90	0.03	0.0000
Household characteristics			
Social economic status			
Poorest	20.96	31.77	0.0000
Poor	21.14	24.50	0.0001
Middle	19.28	20.67	0.0951
Wealth	19.10	14.92	0.0000
Wealthiest	19.52	8.14	0.0000
Location of residence			
Urban	17.37	8.41	0.0000

Appendix 2: Spatial data description and sources

Rainfall data

In the absence of publicly available station rainfall data for Malawi, I used a gridded rainfall dataset called the Climate Hazards Group Infrared Precipitation with Stations (CHIRPS): 2009–2016 Monthly Time Series (Version 2). The CHIRPS dataset, developed by the US Geological Survey (USGS) and the Climate Hazards Group at the University of California, Santa Barbara, couples rainfall data observed from space with more than three decades of data collected at ground stations worldwide. I chose to use the CHIRPS dataset because it has been proven to perform better than ARC (Africa Rainfall Climatology) and TAMSAT (Tropical Application of Meteorology using satellite and other data) in most cases as it has a global coverage at high temporal and spatial resolutions (Dembélé & Zwart, 2016).

Vegetation index data and description

Data for the vegetation index are taken from the MODIS TERRA MOD13Q1.005 250 m dataset, validated version V005 (NASA LP DAAC, 2014), selecting all the images in years 2009–2016 for the tile h18v04. This product provides composite images with a period of 16 days (in which the value of each pixel is the best value considering all 16 days) and two spectral bands, RED (red band, 620–670 nm) and NIR (near infrared band, 841–846 nm). I calculated the average Normalised Differentiated Vegetation Index (NDVI) value for December, January and February prior to the child's month of birth. The NDVI variable was then defined, as the change in average NDVI between two periods, within the primary sampling units. Therefore, there was one NDVI measurement associated with each pregnancy in the primary sampling area.

DHS GPS data and displacement

In order to protect the identities of the DHS respondents, DHS GPS coordinates are displaced (NSO, 2016). The displacement is done randomly accordingly rural points displaced values range between 0 km-5 km of positional error, while urban points displacement values range between 0 km - 2 km.

The child and household data are geographically identified at the primary sampling level in the DHS datasets. Because of the displacement problem, the monthly rainfall shocks indicators

and the average NDVI values are matched to the primary sampling units (PSUs) by calculating 5 km buffers around the displaced cluster centroid. I then merged this data with the DHS GPS data using the latitude and the longitude. Quantum Geographic Information Systems (QGIS) software was used to process the satellite data.

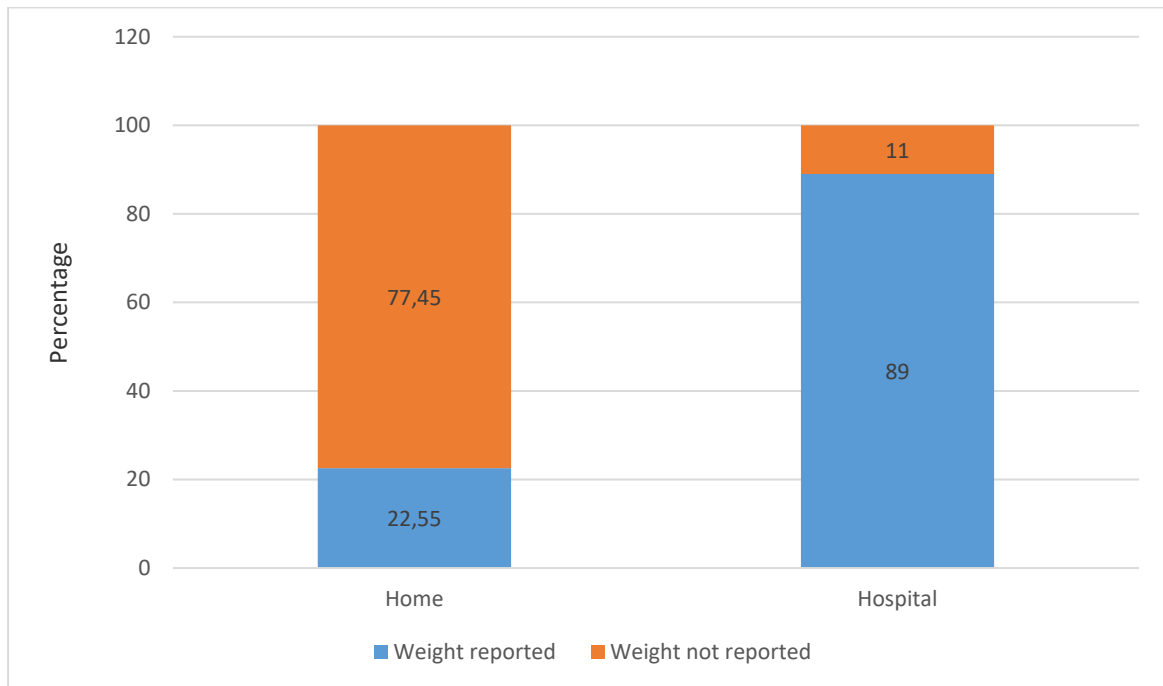## Appendix 3: Institution delivery and probability of reported birthweight

Figure 1: Institution delivery and probability of reported birthweight
Source: Own calculations from 2015/2016 DHS

Appendix 4: First-stage and sample-selection regression models

Table 1: First-stage and sample-selection regression models

Reported birthweight=1	Reported birthweight	ANC utilisation models	
		OLS	**FE**
Instrument			
Institutional delivery	0.595***		
	(0.016)		
Below-average rainfall shocks		0.043***	0.028***
		(0.007)	(0.009)
Maternal age			
Age	0.012***	0.051***	0.049***
	(0.004)	(0.016)	(0.016)
Age2	-0.0002***	-0.001**	-0.001**
	(0.004)	(0.000)	(0.000)
Maternal education			
Primary	0.069***	0.115***	0.059
	(0.013)	(0.042)	(0.044)
Secondary	0.111***	0.163***	0.101*
	(0.015)	(0.051)	(0.053)
Tertiary	0.102***	0.737***	0.633***
	(0.019)	(0.111)	(0.105)
Marital status			
Married	-0.002	0.006	0.011
	(0.008)	(0.030)	(0.029)
Socio-economic status			
Poor	0.014	0.069*	0.078**
	(0.010)	(0.033)	(0.033)
Middle	0.029***	0.053	0.058
	(0.010)	(0.036)	(0.037)
Wealth	0.022**	0.044	0.044
	(0.010)	(0.035)	(0.034)
Wealthiest	0.037***	0.228***	0.235***
	(0.012)	(0.044)	(0.044)

Reported birthweight=1	Reported birthweight	ANC utilisation models	
		OLS	FE
Pregnancy and child characteristics			
Birth order			
2–4th	0.001	-0.138***	0.134***
	(0.010)	(0.035)	(0.036)
Higher order	-0.013	-0.143**	-0.137**
	(0.015)	(0.063)	(0.063)
Twin			
Yes	-0.086***		
	(0.019)		
Sex of the baby			
Boy	0.005		
	(0.006)		
History of Caesarean delivery			
Yes	0.021*	0.209***	0.202***
	(0.012)	(0.045)	(0.044)
History of miscarriage/stillbirth	-0.016	0.038	0.030
Yes	(0.010)	(0.038)	(0.040)
Permission for healthcare seeking			
Big problem	-0.014	-0.045	-0.067
	(0.009)	(0.043)	(0.042)
Residence			
Urban	0.028***	-0.006	0.015
	(0.010)	(0.036)	(0.048)
Seasonality	-	0.036	0.031
		(0.034)	(0.035)
Constant	0.144*	2.766***	2.725***
	(0.083)	(0.216)	(0.220)
Ethnic and religious factors	Yes	Yes	Yes
N	17 223	13 350	13 350
F-statistics on instrument	1409.87***	32.86***	10.37***

Appendix 5: Facility characteristics

Table 1: Comparing facility characteristics between observed and unobserved facilities

Covariates	Observed	Not reported	P-value
Facility type			
Hospital	0.14	0.03	0.0000
Health centre	0.81	0.80	0.9303
Dispensary	0.02	0.04	0.1449
Clinic	0.03	0.13	0.0000
Other	0.003	0.005	0.6696
Managing authority			
Public	0.68	0.64	0.3559
Private	0.32	0.36	0.3559
Facility infrastructure			
Antenatal care area			
ANC services provided 5 days per week	0.93	0.81	0.0000
Private exam room	0.93	0.98	0.3124
ANC guidelines	0.59	0.59	0.9690
Blood pressure apparatus	0.63	0.55	0.0466
Adult weighing scale	0.95	0.93	0.2500
Foetal stethoscope	100	0.96	0.0052
ANC equipment	0.59	0.60	0.8419
Competency			
At least one provider with training in the last two years	0.74	0.55	0.0000
Availability of drugs			
Fansidar	100	0.99	0.2597
Folic iron	0.87	0.87	0.9757
Albendazole	0.99	0.97	0.0188
Toxoid vaccine	0.84	0.80	0.2269

Appendix 6

Table 1: Estimates of ANC indicators by facility observations and DHS women's reports (Women with at least one ANC visit)

Services provided	Direct observation Clusters with facilities that offer the service (%)				Women's DHS reports Clusters with facilities that offer the service (%)			
	All	Urban	Rural	P-val	All	Urban	Rural	P-val
Physical examination								
Blood pressure taken	66.72	70.70	65.54	0.2283	94.04	96.18	93.41	0.1983
Client weight measured	78.63	80.89	77.97	0.4328	99.85	100	99.81	0.5870
Foetal heartbeat assessed	85.90	88.54	85.12	0.2811	100	100	100	-
Screening tests								
Blood sample test for anaemia and syphilis	23.40	34.39	20.15	0.0002	98.69	98.09	98.87	0.4501
Urine tested for protein	11.19	24.20	7.34	0.0000	32.85	43.31	29.76	0.0015
Preventive interventions								
Iron and folic tablets	86.77	88.54	86.25	0.4590	99.85	100	99.81	0.5870
Malaria prophylaxis	77.62	85.99	75.14	0.0041	98.26	97.45	98.49	0.3820
Albendazole	52.76	50.32	53.48	0.4859	66.28	64.33	66.85	0.5574
Offered HIV test	66.57	75.80	63.84	0.0052	100	100	100	-
Counselling								
Nutrition in pregnancy	52.62	64.33	49.15	0.0008	99.85	100	99.81	0.5870
Total sample	**688**	**157**	**531**		**678**	**151**	**527**	