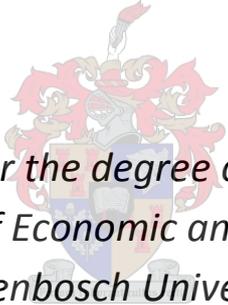


**Healthcare reform priorities for
South Africa:
four essays on the financing,
delivery and user acceptability of
healthcare**

by

Anna Maria Smith



*Dissertation presented for the degree of Doctor of Philosophy in
Economics in the Faculty of Economic and Management Sciences at
Stellenbosch University*

Department of Economics

Stellenbosch University

Private Bag X1, Matieland 7602

South Africa

Supervisor: Professor Ronelle Burger

Co-supervisor: Professor Servaas van der Berg

March 2016

Declaration

By submitting this dissertation electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

This dissertation includes one publication (Chapter 4). The development and writing of this paper was the principal responsibility of myself and a declaration is included in the dissertation indicating the nature and extent of the contributions of co-authors.

The publication of the paper will not infringe upon my right to use the paper in this dissertation or to publish the completed dissertation via Stellenbosch University's SunScholar portal. The article has been reprinted in the dissertation with the permission of the International Union Against Tuberculosis and Lung Disease (Copyright © The Union).

The publication is listed in full below:

Chapter 4:

Smith, A., Burger, R., Claassens, M., Ayles, H., Godfrey-Fausset, P. and Beyers, N. 2016. Health care workers' gender bias in testing could contribute to missed tuberculosis cases in South Africa. *International Journal of Tuberculosis and Lung Disease*. 20(3):350–356. <http://dx.doi.org/10.5588/ijtld.15.0312><http://dx.doi.org/10.5588/ijtld.15.0312>

Copyright ©2016 University of Stellenbosch

All rights reserved

Declarations with respect to co-authoring:

With regard to Chapter 3, the nature and scope of my contribution were as follows:

Nature of contribution	Extent of contribution
Helped formulate research question, literature review, positioning and description of industry context, first attempt at data estimation	35%

The following co-author(s) have contributed to Chapter 3:

Name	Email address	Nature of contribution	Extent of contribution
Rulof Burger	rulof@sun.ac.za	Data estimation and empirical analysis, drafting of data and results sections	65%

Signature of candidate:

Date: 24 September 2015

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 3,
2. no other authors contributed to Chapter 3 besides those specified above, and
3. potential conflicts of interest have been revealed to all interested parties and that the necessary arrangements have been made to use the material in Chapter 3 of this dissertation.

Signature of co-author(s):

Date: 24 September 2015

With regard to Chapter 4, the nature and scope of my contribution were as follows:

Nature of contribution	Extent of contribution
Detailed data analysis, wrote first draft, incorporated all comments received on first and later drafts, expanded version of paper submitted for publication into PhD chapter through a more detailed literature review	65%

The following co-author(s) have contributed to Chapter 4:

Name	Email address	Nature of contribution	Extent of contribution
Ronelle Burger	rburger@sun.ac.za	First attempt at identifying gender patterns in health seeking cascade using dataset, edited and commented on drafts	10%
Mareli Mischa Claassens	mcla@sun.ac.za	Edited and commented on drafts	7.5%
Helen Ayles	Helen@zambart.org.zm	Edited and commented on drafts	5%
Peter Godfrey-Fausset	FaussettP@unaids.org	Edited and commented on drafts	5%
Nulda Beyers	nb@sun.ac.za	Edited and commented on drafts, provided guidance on data analysis and larger TB context	7.5%

Signature of candidate:

Date: 24 September 2015

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 4,

2. no other authors contributed to Chapter 4 besides those specified above, and
3. potential conflicts of interest have been revealed to all interested parties and that the necessary arrangements have been made to use the material in Chapter 4 of this dissertation.

Signature of co-author(s):

Date: 24 September 2015

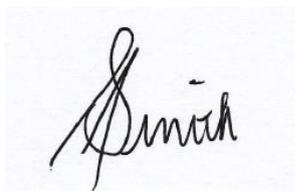
With regard to Chapter 5, the nature and scope of my contribution were as follows:

Nature of contribution	Extent of contribution
Amended questionnaire/instrument used for data collection, managed data collection process, quality control of collected data, data analysis, wrote first draft of chapter, incorporated all edits and comments on various drafts of chapter	90%

The following co-author(s) have contributed to Chapter 5:

Name	Email address	Nature of contribution	Extent of contribution
Vivian Black	vblack@wrhi.ac.za	Developed first version of questionnaire used for data collection for chapter, provided edits and comments on drafts	10%

Signature of candidate:



Date: 24 September 2015

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 5,
2. no other authors contributed to Chapter 5 besides those specified above, and
3. potential conflicts of interest have been revealed to all interested parties and that the necessary arrangements have been made to use the material in Chapter 5 of this dissertation.

Signature of co-author(s):

A handwritten signature in black ink, consisting of a stylized, cursive script that is difficult to decipher. It appears to be a single name or set of initials.

Date: 24 September 2015

Abstract

Despite expenditure levels on healthcare comparable to those of its upper-middle-income country peers, South Africa is achieving health outcomes that are comparable to those of low-income countries.

This dissertation contains four essays on the financing, user acceptability and delivery of healthcare in South Africa. The main contribution of the dissertation is to determine how the user acceptability of healthcare services influences not only health seeking behaviour in South Africa, but also influences the ability of healthcare services to impact health outcomes. Without sufficient focus on user acceptability, the success of the health system will be undermined by creating missed opportunities for the prevention, detection and treatment of disease.

The first essay considers the potential role of private health insurance (medical schemes) in reducing inequality to healthcare access and alleviating the burden from a constrained public healthcare system by providing access to healthcare services of higher user acceptability levels. The analysis indicates that, in the absence of a number of regulatory changes in the market primarily aimed at increasing the affordability of medical schemes, the size of the formal skilled labour market will continue to set the limits of the private health insurance market.

The second essay examines the causal impact of access to private health insurance (medical schemes) on healthcare utilisation and healthcare provider choice by using the exogenous variation in private health insurance coverage induced by the roll-out of the Government Employees Medical Scheme (GEMS). Contrary to most of the findings in the literature, the analysis finds that providing access to healthcare perceived to be of greater user acceptability in South Africa's polarised healthcare market has a large positive effect on total healthcare utilisation. It also increases the likelihood of using private providers and, in particular, private doctors.

In the third essay, the dissertation considers the health seeking behaviour of adults with potential tuberculosis (TB) symptoms (coughed ≥ 2 weeks) in the Western Cape. Only one third of adults indicated they sought help for TB symptoms and only one fourth of those who coughed ≥ 2 weeks reported these symptoms at primary healthcare facilities. Women were less likely than men to be asked for a sputum sample at these facilities, indicating poor adherence by healthcare staff to the well-defined TB testing protocol.

Lastly, the fourth essay explores the causes of late antenatal care access amongst a sample of women in metropolitan Cape Town. More than a quarter of women attended antenatal care late

(≥20 weeks) and, of those who attended late, 48.2% indicated late recognition of pregnancy as the major reason for delayed attendance. While late access was predominantly associated with demand-side factors, late recognition of pregnancy, together with high levels of unplanned pregnancies, point towards issues related to effective access to contraception.

The analysis in the first two essays indicate that there is a demand for healthcare of greater user acceptability, and the last two essays show that this would need to include improved preventative care, enhanced health system effectiveness and better clinical quality monitoring.

Opsomming

Ten spyte van gesondheidsbestedingsvlakke wat vergelykbaar is met dié van ander hoër-middel-inkomste lande, bereik Suid-Afrika gesondheidsuitkomste vergelykbaar met dié van lae-inkomste lande.

Hierdie proefskrif bevat vier opstelle oor die finansiering, gebruikersaanvaarbaarheid en voorsiening van gesondheidsorg in Suid-Afrika. Die hoofbydrae van die proefskrif is om te bepaal hoe die gebruikersaanvaarbaarheid van gesondheidsorg beide gesondheidsoekende gedrag en die vermoë van die stelsel om 'n impak op gesondheidsuitkomste te hê, beïnvloed. Sonder genoegsame fokus op gebruikersaanvaarbaarheid sal die sukses van die gesondheidstelsel ondermyn word deur die verbeuring van geleentede vir die voorkoming, identifikasie en behandeling van siektes.

Die eerste opstel oorweeg die moontlike rol van privaat mediese versekering (mediese fondse) in die vermindering van ongelyke toegang tot gesondheidsorg, sowel as die verligting van die las op die publieke gesondheidstelsel, deur toegang te verskaf tot gesondheidsdienste van hoër gebruikersaanvaarbaarheidsvlakke. Die analise dui daarop dat, in die afwesigheid van 'n aantal regulatoriese veranderinge in die mark hoofsaaklik daarop gemik om die bekostigbaarheid van mediese fondse te verbeter, die grootte van die formele, geskoolde arbeidsmark sal aanhou om die grense van die privaat mediese versekeringsmark te bepaal.

Die tweede opstel ondersoek die kousale impak van toegang tot privaat mediese versekering (mediese fondse) op die gebruik van gesondheidsorg, asook gesondheidsverskafferkeuse deur die benutting van die eksogene variasie in privaat mediese versekeringsdekking wat teweeggebring is deur die uitbreiding van die Regeringswerknemers Mediese Skema (GEMS). In teenstelling met meeste bevindinge in die literatuur, toon die analise dat toegang tot gesondheidsorg van oënskynlik hoër gebruikersaanvaarbaarheid in die konteks van Suid-Afrika se gepolariseerde gesondheidsmark, 'n groot positiewe impak op die benutting van gesondheidsorg het. Dit verhoog ook die waarskynlikheid van die gebruik van privaat gesondheidsverskaffers en, meer spesifiek, privaat dokters.

In die derde opstel stel die proefskrif ondersoek in na die gesondheidsoekende gedrag van volwassenes met moontlike tuberkulose (TB)-simptome (hoes ≥ 2 weke) in die Wes-Kaap. Slegs een derde van volwassenes het aangedui dat hul hulp gesoek het vir moontlike TB-simptome en slegs 'n kwart van dié wat ≥ 2 weke gehoes het, het hierdie simptome aangemeld by 'n primêre gesondheidsorgfasiliteit. Voorts was die neiging dat minder vroue as mans by hierdie fasiliteite vir 'n

sputummonster gevra is, wat dui op op gesondheidspersoneel se swak nakoming van die goed gedefinieerde TB-toetsingsprotokol.

Laastens ondersoek die vierde opstel die oorsake van laat toegang tot voorgeboortesorg in 'n steekproef van vroue in metropolitaanse Kaapstad. Meer as n kwart van die vroue het voorgeboortesorg laat in hul swangerskappe bygewoon (≥ 20 weke) en, van dié wat dit laat bygewoon het, het 48.2% laat bewuswording van swangerskap as die hoofrede vir die vertraging van hul bywoning aangevoer. Terwyl laat bywoning hoofsaaklik geassosieer was met vraagkant-faktore, dui die laat bewuswording van swangerskap, tesame met hoë vlakke van onbeplande swangerskappe in die steekproef, op kwessies rondom effektiewe toegang tot voorbehoedmiddels.

Die analyses in die eerste twee opstelle dui op 'n vraag na gesondheidsorg van hoër gebruikersaanvaarbaarheid en die laaste twee opstelle toon hoe hierdie sorg verbeterde voorkomende gesondheidsorg, hoër stelseleffektiwiteit en beter kliniese gehaltekontrolering sal moet insluit.

Acknowledgements

Although the work contained in this dissertation is my own, I see it as a collaborative project which benefitted from the inputs and support of many people.

I would like to thank Dr. Vivian Black of the Wits Reproductive Health and HIV Institute and Ijeoma Solarin for allowing me to use their research design, previously used in a study in the inner-city of Johannesburg, to explore the causes of late antenatal care access in metropolitan Cape Town. Thank you, Dr. Lungiswa Nkonki, for providing comments on the antenatal care survey instrument. Thank you also to my wonderful fieldwork team who helped with the collection of data for Chapter 5 of the dissertation. Didi Gobile, Lucy Lumphondo, Nomfuzeka Sikota, Christine Abrahams, Mpumi Ketelo and Marché-Lerice Potgieter, you were all excited by the research question and remained committed to the end to help me collect the data.

The financial assistance of the National Research Foundation (NRF) towards my PhD research is hereby acknowledged. All opinions expressed and conclusions arrived at in this dissertation are my own and are not necessarily to be attributed to the NRF.

Thank you to Prof. Ronelle Burger and Prof. Servaas van der Berg for providing funding and mentorship to me during the course of this full-time three year PhD. Without your generosity and teaching this document and process would not have been possible. I am so inspired by your individual commitments to the development economics research profession.

Research on Socio-Economic Policy (ReSEP) in the Economics Department of Stellenbosch University provided an incredibly constructive and supportive environment for my PhD research. I would like to thank some of the following ReSEP colleagues and friends for their enthusiasm for development economics research and their moral support: Marisa Coetzee, Nic Spaul, Laura Rossouw, Carmen Christian and Hendrik van Broekhuizen.

It is difficult to convey the full extent of my gratitude to my parents. Willie en Hannatjie Smith, julle is wonderlike ouers wat my nog altyd onvoorwaardelik ondersteun en liefgehad het in al my keuses en reise. Baie dankie hiervoor.

Lastly, I want to thank the ordinary people of South Africa. This country and its people of contrasts inspire and necessitate deeper thought, creativity and meaning. It asks of us to probe and explore further in order to help solve the complexities which stand in the way of its growth and development.

Table of Contents

Declaration	i
Abstract	vii
Opsomming	ix
Acknowledgements	xi
List of Figures	xv
List of Tables	xvii
List of Abbreviations	xx
List of Definitions	xxii
Chapter 1: Introduction and background	1
1.1 Dissertation structure	9
1.2 Conclusion	12
Chapter 2: Nowhere to grow: exploring the limits of voluntary private health insurance in South Africa	13
2.1 Introduction	13
2.2 The post-apartheid medical schemes landscape	15
2.3 Determinants of health insurance demand	20
2.3.1 Employment and income	21
2.3.2 Risk and health status	24
2.3.3 Consumer choices and healthcare acceptability	26
2.4 Correlates of medical scheme demand	26
2.4.1 Data	26
2.4.2 Methods	33
2.4.3 Results	33
2.5 A middle-market for voluntary private health insurance	41
2.6 Conclusion	44
Appendix A to Chapter 2	47

Appendix B to Chapter 2	52
Chapter 3: Does insurance affect healthcare utilisation and provider choice in a polarised healthcare market? Evidence from a South African natural experiment	55
3.1 Introduction	55
3.2 The relationship between insurance and healthcare utilisation: international evidence	57
3.3 The launch of the Government Employees Medical Scheme (GEMS)	59
3.4 Methodology	61
3.5 Data	63
3.6 Results	66
3.7 Conclusion	71
Appendix to Chapter 3	74
Chapter 4: Gender and TB health seeking and detection: findings from a TB prevalence survey in the Western Cape	75
4.1 Introduction	75
4.2 Literature review	76
4.2.1 Why gender matters for health	77
4.2.2 Gender and TB	79
4.3 Study context	81
4.4 Data	85
4.5 Methods	86
4.6 Results	87
4.7 Discussion	91
4.8 Conclusion	93
Chapter 5: “I didn’t know I was pregnant”: Late access to antenatal care in metropolitan Cape Town	95
5.1 Introduction	95
5.2 The determinants of healthcare utilisation	98

5.3 Maternal mortality in South Africa	101
5.4 The benefits of healthcare that minimises maternal morbidity and mortality	107
5.5 Antenatal care and maternal and infant health outcomes	109
5.6 Factors associated with late antenatal care attendance	113
5.7 Study context	118
5.8 Data	122
5.9 Methods	124
5.10 Results	125
5.10.1 Sample characteristics	125
5.10.2 Pregnancy recognition, confirmation and care seeking	127
5.10.3 Timing of antenatal care attendance and self-reported causes of late access	129
5.10.4 Supply-side factors and timing of attendance	132
5.10.5 Demand-side factors and timing of attendance	138
5.10.6 Regression analysis results	150
5.11 Overview of findings: making sense of results	159
5.12 Conclusion	162
Appendix A to Chapter 5	165
Chapter 6: Conclusion	171
6.1 Chapter 2	173
6.2 Chapter 3	174
6.3 Chapter 4	175
6.4 Chapter 5	176
6.4 What does this mean for the South African health system?	178
References	181
Appendix B to Chapter 5	204

List of Figures

Chapter 2:

Figure 2.1: Comparing trends in medical scheme membership across CMS and survey data	16
Figure 2.2: Medical scheme cover amongst the employed vs. unemployed	17
Figure 2.3: Percentage of black employees with medical scheme cover in the public sector	19
Figure 2.4: Percentage of households that do not have medical scheme cover due to affordability reasons by household expenditure category	23
Figure 2.5: Percentage of total medical scheme members by household expenditure category	23
Figure 2.6: Kernel density function of age distribution of all medical scheme members for 2009-2011	30
Figure 2.7: Kernel density function of age distribution of white medical scheme members for 2009-2011	31
Figure 2.8: Kernel density function of age distribution of black medical scheme members for 2009-2011	31

Appendix A to Chapter 2:

Figures 2.A1: Kernel density function of age distribution of all medical scheme members for two time periods: 2002-2004, 2005-2008	49
Figures 2.A2: Kernel density function of age distribution of black medical scheme members for two time periods: 2002-2004, 2005-2008	50
Figures 2.A3: Kernel density function of age distribution of white medical scheme members for two time periods: 2002-2004, 2005-2008	50

Chapter 3:

Figure 3.1: Total lives (members and beneficiaries) covered by GEMS, 2006-2013	60
Figure 3.2: Growth in medical scheme members and beneficiaries, 2002-2011	65
Figure 3.3: Share of working age individuals with medical scheme coverage, 2003-2008	66

Chapter 4:

Figure 4.1: Reported cases of all TB (pulmonary, extra-pulmonary, drug-resistant, extremely	
---	--

drug-resistant per 100,000 of the population as obtained from TB registers	82
Figure 4.2 TB cure rate for new smear-positive cases as obtained from TB registers	83
Figure 4.3 Health seeking and TB diagnostic cascade of all adults who coughed ≥ 2 weeks (by gender, 11,297 men and 18,720 women)	88
Figure 4.4 Health seeking and TB diagnostic cascade of culture-positive adults (by gender, 333 men and 369 women)	90
Chapter 5:	
Figure 5.1: Maternal mortality ratio of various countries (2010) relative to per capita government expenditure (PPP in US\$ for 2011)	104
Figure 5.2: Recent institutional maternal mortality ratios (iMMRs)	106
Figure 5.3: South African national rate of antenatal care attendance at public health clinics before 20 weeks gestation	118
Figure 5.4: Western Cape rate of antenatal care attendance before 20 weeks gestation	119
Figure 5.5: Antenatal care seeking cascade for all women	128
Figure 5.6: Timing of antenatal care attendance	129
Figure 5.7: Travel time to clinic for those who attended antenatal care	133
Figure 5.8: Cost of one-way trip to clinic for those who attended antenatal care and used public transport to the clinic	134
Figure 5.9: Services that were provided by clinic at booking (first screening) visit	135
Figure 5.10: Distribution of age of booked by early (<5 months) vs. late antenatal care attendance (≥ 5 months)	141
Figure 5.11: Percentage of observations (total sample, N=221) by household income response category	142
Figure 5.12: Proportion of women who accessed antenatal care late (≥ 3 months) by asset index quintile	146
Figure 5.13: Proportion of women who indicated they have an unplanned pregnancy about which they are unhappy (unwanted pregnancies) by asset index quintile	149

List of Tables

Chapter 2:

Table 2.1: Output of linear probability model using LFS and QLFS data (2000-2011)	35
Table 2.2: Output of linear probability model using GHS (2009-2011)	38

Appendix A to Chapter 2:

Table 2.A1: Number of medical scheme main members and members from the combined Labour Force Surveys (LFS) and Quarterly Labour Force Surveys (QLFS), the Council for Medical Schemes and the General Household Surveys (GHS)	47
Table 2.A2: Total medical scheme coverage by race	47
Table 2.A3: Black employment and medical scheme cover	48
Table 2.A4: White employment and medical scheme cover	48
Table 2.A5: Reported chronic disease and illness (or injury) amongst those with medical scheme cover and those without	49
Table 2.A6: Output of linear probability model using GHS: 2002-2005, 2002-2011	51

Chapter 3:

Table 3.1: Impact of insurance on utilisation, 2002-2008 (LFS/QLFS, GHS)	67
Table 3.2: First-stage estimates of likelihood of being insured (NIDS 2008-2012)	68
Table 3.3: OLS estimates of effect of health insurance on choice of health facility	70
Table 3.4 2SLS estimates of effect of health insurance on choice of facility	71

Appendix A to Chapter 3:

Table 3.A1: The effect of GEMS on health insurance, by sector	74
---	----

Chapter 4:

Table 4.1: Provincial distribution of TB sample	84
Table 4.2 Gender distribution by province of adults that reported being ill with TB	84
Table 4.3: Health seeking behaviour and TB diagnostic care received by adults who coughed ≥ 2 weeks (by gender)	87
Table 4.4: Provider choice by gender for adults who coughed ≥ 2 weeks and sought care	87
Table 4.5 Characteristics of adults (by gender) in prevalence survey who coughed ≥ 2 weeks	

and sought care at primary healthcare facility	89
Table 4.6: Health seeking behaviour and TB diagnostic care received by culture-positive TB adults in prevalence survey (by gender)	90
Chapter 5:	
Table 5.1: Maternal and child health indicators for study sub-districts	121
Table 5.2: Socio-demographic characteristics of total sample	126
Table 5.3: How pregnancy was confirmed	128
Table 5.4: General reasons for late antenatal care attendance	130
Table 5.5: Personal factors that could have enabled earlier attendance	131
Table 5.6: Clinic factors that could have enabled earlier attendance	132
Table 5.7: Services not offered by clinic at booking visit (early vs. late attenders)	136
Table 5.8: Reported weak performance on clinic quality indicators by early vs. late antenatal attendance	138
Table 5.9: Socio-economic characteristics and health behaviour by government definition of early (<5 months) vs. late (≥5 months) antenatal care attendance	140
Table 5.10: Ownership of durable assets from survey vs. ownership of selected same assets from NIDS	143
Table 5.11: Socio-economic characteristics and health behaviour by alternative definitions of early (<3 months) vs. late (≥3 months) antenatal care attendance	144
Table 5.12: Socio-economic and health characteristics/behaviour of respondents by unwanted pregnancies vs. wanted pregnancies	147
Table 5.13: Variables included in regression analysis – expected sign and rationale	151
Table 5.14: Regression results for late access (narrow definition, ≥5 months) to antenatal care	153
Table 5.15: Regression results for late access (broad definition, ≥3 months) to antenatal care	155
Table 5.16: Regression results for unwanted pregnancies	157
Appendix A to Chapter 5:	
Table 5.A1: Variables used in estimation of asset index and weights/factor loadings obtained from MCA	165

Table 5.A2: Regression results for late access (narrow definition) by unwanted vs. wanted pregnancies	166
Table 5.A3: Regression results for late access (narrow definition) by first pregnancy vs. second or later pregnancies	167
Table 5.A4: Regression results for late access (broad definition) by unwanted vs. wanted pregnancies	168
Table 5.A5: Regression results for late access (broad definition) by first pregnancy vs. second or later pregnancy	169
Table 5.A6: Regression results for timing of antenatal care access by pregnancy month	170

List of Abbreviations

AIDS	Acquired immune deficiency syndrome
ANC	African National Congress
ART	Anti-retroviral therapy
BANC	Basic Antenatal Care
CHC	Community health centre
CMS	Council for Medical Schemes
DHIS	District Health Information System
GEMS	Government Employees Medical Scheme
GHS	General Household Survey
GP	General practitioner
HAART	Highly active antiretroviral therapy
HDACC	Health Data Advisory and Co-ordination Committee
HIV	Human immunodeficiency virus
iMMR	Institutional maternal mortality ratio
IV	Instrumental variable
LFS	Labour Force Survey
LIMS	Low-income Medical Schemes
LPM	Linear probability model
MCA	Multiple correspondence analysis
MMR	Maternal mortality ratio
MOU	Maternity obstetric unit
NCCEMD	National Committee for Confidential Enquiries into Maternal Deaths
NHI	National Health Insurance
NIDS	National Income Dynamics Survey
NPRI s	Non-pregnancy related infections
PHC	Primary healthcare

PMBs	Prescribed minimum benefits
PSU	Primary sampling unit
QLFS	Quarterly Labour Force Survey
REF	Risk equalisation fund
TB	Tuberculosis
UHC	Universal health coverage
UN	United Nations
WCDOH	Western Cape Department of Health
WHO	World Health Organization
ZAMSTAR	Zambia, South Africa Tuberculosis and AIDS Reduction (ZAMSTAR) trial

List of Definitions

- Early access:* Accessing antenatal care <20 weeks of pregnancy as per the definition used by the South African District Health Information System (DHIS) and Department of Health to monitor timing of antenatal care access.
- Late access:* Accessing antenatal care ≥ 20 weeks of pregnancy as implied by the definition of early antenatal care access of the South African District Health Information System (DHIS) and Department of Health.
- Late access narrow definition:* Same as late access defined above.
- Late access broad definition:* Accessing antenatal care ≥ 12 weeks of pregnancy (at or after the end of the first trimester).
- TB symptomatic:* A cough with a duration of ≥ 2 weeks or coughing ≥ 2 weeks.
- Unwanted pregnancy:* An unplanned pregnancy about which a woman was unhappy about when she found out she was pregnant.
- Wanted pregnancy:* A planned pregnancy, or an unplanned pregnancy about which a woman was happy about when she found out she was pregnant.

Chapter 1

Introduction and background

Inequality in health outcomes can be considered both a cause and consequence of income inequality. Health outcomes are influenced by many inputs outside the immediate control of the health system, including public goods such as sanitation and housing (Jack and Lewis, 2009). Long-term improvements in health outcomes are, in fact, ascribed by multiple authors (e.g. Jack and Lewis, 2009; Lleras-Muney, 2005 and McKinley and McKinley, 1997) to improvements in lifestyle (including education) that is a consequence of the growth and development of countries over time, rather than to direct expenditure on health and the growth and expansion of health systems.

Health or, more specifically, poor health, can also impact income. It is well-documented that large, unexpected health events can lead individuals to enter states of poverty due to the expenditures required to purchase access to healthcare and the other indirect costs associated with these events (e.g. Van Doorslaer et al., 2001; Wagstaff and van Doorslaer, 2003; Whitehead, Dahlgren and Evans, 2001; Xu et al., 2003), especially in the absence of well-functioning health financing systems. The catastrophic health expenditure¹ associated with such health events can also cause individuals to remain trapped in poverty.

A well-functioning and accessible health system should be able to actively influence health outcomes and address and counteract inequality in health outcomes, thereby also leading to a reduction in income inequality and poverty over the longer term. Ultimately, in interacting with the health system, individuals should have equality of opportunity to access and utilise the system, irrespective of their individual life circumstances. Clearly this means that the system cannot simply offer the same to everyone because vulnerable sub-groups may have complicated lives that obstruct their ability to utilise the health system when needed. Similarly, there may be social norms that may inhibit health seeking amongst certain groups. This means that to be fair, the system may have to compensate for these obstacles that could constrain health seeking behaviour amongst vulnerable groups, e.g. women, individuals in poverty or even the youth.

The preceding discussion should make it clear that a definition of vertical equity in healthcare applies. This definition of equity applies throughout the dissertation. The alternative would have

¹ Catastrophic health expenditure refers to health expenditure in excess of a pre-defined income threshold that may have an impoverishing effect on households. Various thresholds have been proposed. Two widely used thresholds is healthcare expenditure that exceeds 10% of total household expenditure and healthcare expenditure that exceeds 40% of total household non-food expenditure (O'Donell, 2008 as cited in Burger et al., 2012; see also the seminal paper by Xu et al., 2003).

been to employ a definition of horizontal equity. Culyer (2001:276) defines the distinction between vertical and horizontal equity as follows: “Horizontal equity requires the like treatment of like individuals and vertical equity requires the unlike treatment of unlike individuals, in proportion to the differences between them.”

Even if health services, however, seem to be accessible on the surface, if the system does not function at optimal levels of effectiveness and if there is wastage present, the ability of the system to impact health outcomes will be severely compromised. For instance, even if clinics are close to users and have convenient opening hours, but are missing opportunities to diagnose and treat diseases, then health outcomes will not improve. Any expenditure on such a system will lead to sub-optimal health outcomes relative to what could have been achieved. It is then possible to ask whether the same expenditure on other development goods (e.g. sanitation, housing and nutrition) could have led to an equivalent or even larger improvement in health outcomes².

There is growing recognition that the provision of physical access to healthcare is unlikely to improve health outcomes if the quality of the provided healthcare is inadequate (Das, Hammer and Leonard, 2008; Das and Hammer, 2014). Das and Hammer (2014) argue that globally, and in most developing countries, there is a sufficient supply of healthcare services, i.e. there are enough buildings, health equipment and even health staff. Individuals in many developing countries tend to also have an average number of health visits comparable to individuals in developed countries (Das and Hammer, 2014). In this context, health outcomes that remain weak or sub-optimal can therefore be ascribed to poor quality healthcare services. It is argued that, ultimately, it is the quality of the “clinical encounter” that matters for health outcomes – [this] “has to do with the accuracy of the advice and it is this accuracy that represents the true value added of the provider” (Das and Hammer, 2014).

A polarised health system:

This dissertation explores some of the above questions and, specifically, the failure of the South African health system to lead to large improvements in health outcomes. Before providing more

² If the focus is placed more on poverty or inequality reduction rather than only health outcomes, a second and potentially more important question would be whether the same expenditure on different, non-health related development “goods” such as education and direct income transfers could achieve a greater reduction in income inequality and poverty. This question was recently explicitly raised by the development economist, Jeffrey Hammer, in the context of the “Economists’ Declaration on Universal Health Coverage”, a petition requesting support for universal health coverage from eminent economists globally. Hammer (2015) argued as follows: “Whether health care is particularly important for poor...must be evaluated against everything else governments might do to rectify an unfair distribution of income. Health care is not an obvious choice in comparison to food, for example, or unconditional cash transfers.” This dissertation does not, however, explicitly deal with the second question, but it is important to keep this in mind in thinking about the ability of healthcare to reduce and prevent poverty and inequality.

background on the possible reasons why health outcomes may not be improving to the degree required I first provide background context on the nature of the South African health system, including the size, scope and financing of the public and private sectors.

South Africa's health expenditure levels compares well to those of its upper-middle-income peers. In 2013, total health expenditure constituted 8.9% of gross domestic product (GDP) (WHO, 2015a). This is higher than the average share of 6.3% of total health expenditure as percentage of GDP amongst upper-middle-income countries, slightly below Brazil's 9.7% of total health expenditure as share of GDP and exceeds China's of 5.6% (World Bank Development Indicators, 2015).

In 2013, total government expenditure on health was equivalent to 14.0% of total government expenditure (WHO, 2015a). Government expenditure totalled 48.4% of total health expenditure in the same year. Private health expenditure therefore totalled 51.5%, with private health insurance accounting for 81.1% of total private health expenditure. It is estimated that only 7.1% of total health expenditure was paid for on an out-of-pocket basis (WHO, 2015a).

The majority of government health expenditure goes towards funding a non-contributory, tax-funded public health system consisting of primary, secondary and tertiary levels. General taxes, which are used to fund the system, are allocated to nine provinces and the National Department of Health (Van den Heever, 2012).

User fees for primary healthcare services were abolished in the mid-1990's (McLeod et al., 2007). A user fee system, subject to a means test, still applies to public healthcare services at secondary and tertiary levels. However, as McLeod and Grobler (2008) note, "the exemption policy is liberally applied and so for the unemployed and very low income earners, care is provided virtually free at point of service".

Although there is variation on year-on-year basis, medical scheme (private health insurance) coverage has stayed relatively stable, providing cover to between 16.0% and 17.0% of the population who mainly use private healthcare services. By the end of 2013, 8.8 million South Africans (Council for Medical Schemes, 2014) or approximately only 17.0³ of the population were medical scheme members.

The 83.0% of South Africans (approximately 44 million) not covered by a medical scheme mainly utilise public healthcare services, but data from 2007 indicates 28.8% first sought care at a private

³ Council for Medical Schemes total calculated as percentage of estimated total South African population of 52.98 million as per Statistics South Africa Mid-Year Population Estimates (2013).

facility when they were ill (Van Eeden, 2009). Even amongst the poorest quintile of households about 20.0% access private healthcare services (Burger et al., 2012).

Health outcomes and the disease burden:

Studies describing the nature of the South African health system conclude that despite South Africa's status as an upper-middle-income country, it is producing worse health outcomes than many lower income countries (e.g. Van den Heever, 2012; DBSA, 2008).

Any analysis of healthcare in South Africa should also be cognisant of the large disease burden. South Africa's health system is confronted with a so-called "quadruple" burden of disease, with high prevalence and incidence of communicable diseases (e.g. tuberculosis, or TB, and HIV/AIDS), an increasing burden of non-communicable diseases (e.g. diabetes and cardiovascular disease), high prevalence of injuries of which many are due to high levels of interpersonal violence within an unequal and conflicted society, and maternal and child health problems (Mayosi et al., 2009).

The large financing gap between the public and private health sectors in South Africa is frequently identified as a possible cause of the under-performance of the public health sector. According to Mills et al. (2013: 133), "South Africa has the largest share in the world of total health-care expenditure funded through private insurance (44%), yet only 16% of the population benefit from these resources" (more information of the financing of the system is provided below).

The consistent under-performance in critical public health areas, e.g. communicable diseases and maternal and child health, relative to peer countries cannot, however, simply be explained away by the high disease burden (specifically HIV) or financing inequity between the public and private health sectors. Rather, it more likely indicates ineffectiveness or wastage in the public health system. While this is an issue that has not yet been well-explored in the South African context, there exists some evidence that alludes to the presence of x-inefficiency⁴ in the public health sector once the burden of disease and available resources have been controlled for. Christian and Crisp (2012) cite several examples of a lack of leadership and decision-making power in the public health system that is likely to lead to x-inefficiency. Variation in health outcomes between public health facilities on provincial and district level also provides an indication of x-inefficiency. According to Engelbrecht and Crisp (2010: 201) "there are numerous examples of relatively poorly-resourced districts that have better

⁴ Christian and Crisp (2012: 726) define this as "an open-ended concept which describes the effectiveness with which a set of inputs can produce outputs" first used by Leibenstein in 1979. While it is closely related to the concept of technical efficiency, it differs from it in that the source of inefficiency is explicitly identified as "intrinsic to the nature of human behaviour", e.g. management and decision-making in organisations (Christian and Crisp, 2012: 727).

outcomes than very well-resourced districts". Palmer et al. (2002) finds that a sample of low-cost private sector clinics in South Africa is able to provide higher quality services, as measured by client perceptions, at similar cost to public sector clinics.

Missed opportunities - diagnosing sources of health system under-performance in South Africa:

There is a growing body of evidence from the medical and public health literature that points towards under-performance and system ineffectiveness in the delivery of public healthcare services of all levels in South Africa. This includes primary healthcare services as provided by clinics and secondary and tertiary healthcare services as provided by different types of hospitals.

At a primary healthcare level, there exists evidence of sub-optimal performance in critical public healthcare areas such as maternal health and HIV and TB care.

In the delivery of antenatal care, Solarin and Black (2013) found that almost 50.0% of women who sought antenatal care at primary healthcare facilities in inner-city Johannesburg were not screened (had a "booking" visit) or seen by a nurse at their first visit to the clinic. Of this group, 39.2% experienced a delayed booking visit after their first clinic visit because clinic staff told them to return more than a month later, leading to an average delay of 3 months in having a "booking" visit at the antenatal clinic after being told to return later. This occurred despite the fact that official health policy requires these clinics to see women on the day they first present at the clinic.

When disease is diagnosed, the health system often responds slowly with the initiation of treatment. Amongst a sample of HIV-positive pregnant women in Johannesburg, Myer et al. (2012) found an average delay of three weeks between screening for HIV and ART initiation. This delay was found to not improve the health outcomes of these women before and after giving birth.

Similar delays in treatment initiation, and even the loss of patients from the system, have been documented in the TB disease context. A comparison of the sputum register⁵ with the TB treatment register for 122 primary healthcare facilities in five provinces of South Africa found a mean initial loss to follow up rate of 25.0% (Claassens et al., 2014). This implies that, on average, 25.0% of individuals diagnosed with TB in these facilities were not initiated on treatment. They were lost to the system. A significant association was found between the initial loss to follow up rate and the turnaround time of sputum test results (time between sample being taken and results received by the facility). The study found that "the main determinant of the ILF [initial loss to follow up] was the proportion of

⁵ Document/electronic registry containing personal information on all individuals who provided sputum samples for TB testing.

sputum results returned to the facility in 48 hours” (Claassens et al., 2015: 605). Across facilities, a mean of only 42.8% of results were received back at the facility within 48 hours.

South Africa’s maternal and child healthcare also provide evidence on sub-optimal system performance at hospital level. A recent audit report concluded that hospital performance and the quality of clinical care played a significant and increasingly important role in explaining maternal deaths in South Africa. The report found a 25.0% increase in maternal deaths due obstetric haemorrhage compared to the earlier assessment period and identified poor clinical assessment, delays in referral, not following standard protocols and not responding to abnormalities in the monitoring of patients as the most common avoidable hospital-related factors (Moodley, 2014). A series of quality-of-care audits in hospitals with high and increasing levels of perinatal deaths identified a number of hospital-related factors contributing to these deaths. These factors included the lack of antenatal steroids (which speaks to pharmaceutical stock management), insufficient nursing staff, fetal distress not being monitored and, in cases where the fetus was monitored, poor progress in labour with incorrect interpretation of the partogram⁶ (Allanson and Pattinson, 2015).

At the same time, the users of healthcare services, especially public healthcare services, in South Africa have certain socio-economic characteristics typically associated with late or low health seeking behaviour. Social exclusion and social norms that constrain health seeking may cause individuals to disengage from the health system, or may lead them to never engage in the first place. These socio-economic characteristics include lower education levels, or educations of inferior quality, compared to their counterparts who primarily utilise private healthcare services. Lower income levels are another such characteristic which is likely to influence both the direct and opportunity costs of seeking healthcare. Individuals with these types of socio-economic characteristics are also likely to face greater exposure to harsh environmental factors due to poor housing and sanitation and a greater exposure to diseases in their everyday living conditions. Due to greater exposure to illness and other risk factors, there will be greater need for healthcare amongst these vulnerable groups.

The delivery of healthcare services in a low access context, or the delivery of services of lower user acceptability⁷ levels, is likely to influence both healthcare seeking behaviour (on the demand-side)

⁶ A graphic representation of the most important maternal and fetal health data collected during the labour process which can indicate deviations from labour process norms (WHO Reproductive Health Library, 2014).

⁷ A detailed discussion on the term acceptability within a healthcare access framework context is provided in Chapter 5 (Section 5.2) of the thesis. The term user acceptability is used to refer to both the quality dimension of healthcare access as identified by Goddard and Smith (2001) and “softer”, cultural perception issues which are likely to influence the behaviour of both clients and healthcare providers, as included under the McIntyre, Thiede and Birch (2009) access dimension of acceptability.

and the ability of health system to detect health conditions and thereby influence health outcomes (supply-side). This is illustrated by the high rate of loss to follow up in South Africa's TB context and the importance of delays in obtaining sputum test results in explaining the high loss to follow-up rate (Botha et al., 2008; Claassens et al., 2014). While much of the high loss to follow up rate can be explained or, at the least, partly explained by supply-side factors, there also exists evidence of the influence of demand-side factors. Some of the patients who do not return to clinics to obtain their TB sputum test results and initiate treatment, those classified as being part of the loss to follow up rate, provide incomplete address details or move around, pointing towards potential socio-economic vulnerability (Botha et al., 2008).

It is important to understand how demand- and supply-side factors, and the interaction of these factors, contribute to the weak health outcomes achieved in South Africa.

The health system reform agenda attempts to address under-performance:

While the private healthcare sector is also characterised by the presence of inefficiencies and ineffectiveness, and there are concerns about anti-competitive behaviour and market structure problems (Halse et al., 2012), the focus in the discussion below and in this dissertation falls mainly on under-performance in the public sector. The public sector is the part of the health system that serves the majority of South Africans and is tasked with having to manage and treat the brunt of the South African disease burden. It is also the part of the system that is best positioned to deal with public health concerns such as maternal and child health and communicable diseases that have large externalities associated with them.

An official government proposal has been made to move South Africa to a universal health coverage (UHC) system through a National Health Insurance (NHI) Fund⁸ (see Department of Health, 2011a). However, as Van den Heever (2012: S5) points out, from a financing perspective, South Africa effectively already has a UHC system: "Given this configuration [referring to funding of public and private health systems], South Africa technically complies with the goal of universal coverage as a

⁸ The paper (Department of Health, 2011a) criticised the current health financing system mainly based on equity concerns and increasing costs experienced in the private health insurance market. It set out the principles and objectives of the proposed NHI as a solution to inequity in healthcare expenditure between the public and private sectors and provided highlights of the planned key components (both infrastructure and general operations) of the system. While it provided high-level cost estimates for the new system, it did not provide any detail on costing assumptions. It also set out ambitious timeline for the implementation of NHI. At the time of writing this thesis, only the Green Paper had been released. The White Paper to the NHI proposal was released in December 2015 (Republic of South Africa, 2015). Although this thesis recognises its release, it does not deal with the content and implications of the document.

comprehensive package of health services is available on a pre-paid basis either through the public sector or regulated health insurance [medical schemes].”

Rather, at its core, the NHI proposal contained in the official government “green paper” or NHI discussion document released in August 2011, should be viewed as an attempt to rectify some of the performance issues in the current South African public health system. While the proposal focuses on financing and is an attempt to increase total (public) health expenditure in South Africa by harnessing private funding, many of the supporting proposals focus on the re-engineering of the way that public health services are provided⁹. Although not always explicitly positioned as such, these proposals can be viewed as attempts to improve the quality of public healthcare services and explicitly address the polarisation (both in financing and perceived user acceptability) between the public and private health sectors.

Following the release of the 2011 green paper, the Department of Health and National Treasury were required to develop more detailed proposals on NHI. Both these documents were expected to be finalised early-2013, but at the time of writing neither of these documents have been publically released.

The majority of the discussion and context on health services in the NHI green paper was focused on the inequity in funding between the public and private healthcare markets. The document did not, however, directly deal with sources of ineffectiveness in the public health sector or how health seeking behaviour and the interaction of health seeking behaviour and choices with the nature of public healthcare delivery may contribute to weak health outcomes in South Africa. From a reading of the green paper, it is clear that much of the under-performance of the public healthcare system is attributed to differences in both the scope and nature of funding between the public and private sectors.

An alternative perspective on the health system’s poor performance is offered by recent work by Andrews, Pritchett and Woolcock (2013). States (governmental) or sub-components of states such as health systems may get stuck in conditions of low functioning or capability traps (Andrews, Pritchett and Woolcock, 2013). One of the features of a capability trap is that form or appearance takes primacy over function. Low capability traps are often reached through a process of “isomorphic mimicry”, where government systems tend to adopt policies from the systems of other higher functioning states (Andrews, Pritchett and Woolcock, 2013). While on paper it may seem that these

⁹ The main focus was on changing the existing primary healthcare system to deliver services through district-based clinical support specialist teams, school-based primary healthcare services and “municipal ward-based primary healthcare agents” (i.e. community healthcare workers) (Department of Health, 2011a).

systems comply with international best practice, the borrowed policies do not filter through to practice and implementation, often because the policies themselves were never suitable to the conditions in which they were implemented (Andrews, Pritchett and Woolcock, 2013)¹⁰.

This dissertation provides evidence of how different dimensions of access to healthcare services can influence health outcomes. The essays contained in this dissertation, in particular, emphasise financing (mainly private services) and user acceptability of healthcare (both public and private services). The dissertation also considers how demand-side factors related to socio-economic vulnerability contribute to sub-optimal health outcomes by interacting with a constrained supply-side in the delivery of healthcare services. Much of the focus of the dissertation thus falls on the nexus between the demand- and supply-sides of the South African health system.

1.1 Dissertation structure

This discussion and the evidence on system under-performance from the larger medical or public health literature implies that a detailed evidence-based is required for, on the one hand, sources of health system ineffectiveness (delivery) and, on the other hand, a much deeper understanding of health system users' experiences of the system, their socio-economic contexts and whether and how this influences health seeking behaviour.

The main contribution of this dissertation is, therefore, to consider how access to health services, and acceptability, influences not only health seeking behaviour, but also manifests in the ability of healthcare services to impact health outcomes. This is of critical importance in embarking on any major healthcare reform process as is the case in South Africa.

Healthcare access is a broad concept that encompasses affordability, availability and acceptability of healthcare services (McIntyre, Birch and Thiede, 2009)¹¹. While the direct costs for users of public healthcare services should not be a major deterrent, indirect costs may influence health seeking behaviour. Costs do not, however, seem to be a major constraint to public healthcare in South Africa. The physical supply of services (availability) has also been significantly broadened in post-apartheid South Africa with a large-scale expansion of the clinic network and a relative increase in expenditure on clinics as the front-line health facility (Burger et al., 2012). The dissertation therefore seeks to demonstrate that user acceptability now forms the final frontier in establishing effective

¹⁰ Andrews, Pritchett and Woolcock (2013) propose that low capability traps can be escaped through a gradual process of "problem-driven iterative adaptation". This process does not entail large reforms adopted from other countries but rather encourages and rewards incremental experimentation, with well-defined "active learning mechanisms" and "iterative feedback loops" (Andrews, Pritchett and Woolcock, 2013: 239).

¹¹ These concepts are defined in greater detail in Chapter 5.

access to healthcare in South Africa. This focus on user acceptability is the Leitmotif through the various chapters of the dissertation and is considered in the contexts of private health financing, TB and antenatal care.

More information on the focus of each chapter is provided below.

Chapter 2:

Against the backdrop of the proposed National Health Insurance (NHI) scheme, this chapter considers the potential role of private health insurance, in the form of medical schemes, in improving equity in healthcare access in South Africa. The existence of medical schemes as a healthcare funding mechanism allow consumers greater choice in healthcare services and, at the least, allows for the bypassing of rationing mechanisms in the public sector.

The chapter first considers compositional changes in the medical schemes landscape in post-apartheid South African. Using pooled data from different household surveys, I then consider the associations between medical scheme membership and various socio-economic and demographic factors using linear probability models (LPMs). The role of user acceptability of healthcare services and physical access to healthcare services as potential determinants of health insurance membership are explicitly explored in the analysis. I also consider the growth opportunities for medical scheme membership in South Africa to alleviate the burden on a constrained public healthcare system and potentially assist in reducing inequality in healthcare access.

Chapter 3:

A reading of the current causal empirical literature suggests that health insurance only affects healthcare utilisation in countries where the healthcare system is highly polarised. I investigate this hypothesis in the context of South Africa, where apartheid era policies left the country with a highly polarised healthcare system with a stark division between the public and private sectors. In order to estimate the causal effect of health insurance on health seeking behaviour, I exploit the exogenous variation in medical scheme coverage induced by the establishment of the Government Employees Medical Scheme (GEMS) in 2006 and the gradual roll-out of membership that occurred from 2007 onwards.

Two datasets are used to test, firstly, the effect of the initial implementation of this policy in 2007 and, secondly, the effect of the continued roll-out between 2008 and 2012. The identification strategy uses aspects of difference-in-difference and instrumental variable estimators to identify the

causal effect of health insurance on health seeking behaviour. Using this approach, I explore the effect of access to private health financing on total healthcare utilisation and provider choice.

Chapter 4:

In this chapter, I consider the health seeking behaviour of TB-symptomatic adults (coughed ≥ 2 weeks) from a gender perspective in the Western Cape. The detailed, step-by-step analysis of health seeking behaviour and choices, starting from the identification of TB-symptomatic adults, allow for the identification of missed opportunities in the larger context of TB detection and treatment.

Two datasets are used: firstly, a pooled nationally representative household survey from 2002 to 2011 to provide a larger context to TB prevalence in South Africa and, secondly, a TB prevalence survey which was conducted in 2010 as part of a larger TB intervention in eight high TB-prevalent communities in the Western Cape.

The chapter starts by reviewing literature on gender and health, and, gender and TB, respectively, and then moves to provide more information on the study context and methods. Next, I analyse the TB prevalence survey by using bivariate analysis to consider gender patterns in the TB health seeking cascade. The socio-economic correlates of TB symptomatic adults are also considered. The results of the data analysis are discussed and interpreted by referring to how the results relate to other study findings. The chapter is concluded by briefly referring to the policy implications of the findings.

Chapter 5:

Early access to appropriate antenatal care can ensure the detection and treatment of diseases and health conditions that impact maternal and child health. Although South Africa has high levels of antenatal care coverage and deliveries in healthcare facilities and is almost achieving the minimum number of antenatal care visits recommended by the World Health Organization (WHO), a large proportion of women access antenatal care late (≥ 20 weeks/5 months) in their pregnancies. Early antenatal care attendance allows for the optimal diagnosis and treatment of HIV, a major cause of maternal death in South Africa. It also enables the diagnosis and treatment of any other pregnancy risk conditions (e.g. high blood pressure and anaemia) and diagnosis of possible fetal health conditions.

To explore the issue of late antenatal care access, I used the methodological approach of Solarin and Black (2013) and conducted a cross-sectional survey at four public sector labour wards in metropolitan Cape Town between October and November 2014. A total of 221 women were interviewed.

The survey captured self-reported timing of first antenatal care access. Using univariate, bivariate and multivariate analysis in the form of LPMs, I explore late antenatal care access relative to various socio-economic, demographic, health and health system factors. In particular, I explore late antenatal care access relative to the three dimensions of healthcare access, namely availability, affordability and acceptability.

1.2 Conclusion

To summarise, the aim and contribution of this dissertation is to examine issues related to the financing, user acceptability and delivery of healthcare services in South Africa in the context of the large-scale proposed health system reforms. It first considers South Africa's private health insurance system (the medical schemes market) and how and whether this market has the potential to reduce inequality in access to healthcare services of perceived higher user acceptability levels. It then moves to consider the causal impact of the extension GEMS to South African government workers on healthcare utilisation and provider choice – it essentially focuses on the demand response when more people are provided with access to private healthcare services.

Next, in the context of South Africa's high TB burden, the dissertation applies a gender lens to health seeking behaviour and the detection of TB in primary healthcare facilities and also considers missed opportunities in the diagnosis of TB. Lastly, it measures and considers the self-reported causes and correlates of late antenatal care attendance in metropolitan Cape Town in the context of South Africa's high maternal mortality levels.

The findings from these four chapters or essays serve to provide a better diagnosis of how the ability of the health system to impact health outcomes is compromised by, on the one hand, the interaction between social exclusion and social norms promoting less than ideal health seeking behaviour (the demand-side) and, on the other hand, health system ineffectiveness (supply-side). Poverty and inequality levels, which contribute to sub-optimal health seeking and social exclusion, are likely to change slowly in South Africa. This means that the health system has to operate as effectively as possible if it is to improve and impact health outcomes. Increases in system effectiveness are likely to be experienced as increased or greater user acceptability which, in turn, is likely to influence health seeking behaviour and lead to improved health outcomes.

Chapter 2

Nowhere to grow: exploring the limits of voluntary private health insurance in South Africa

2.1 Introduction

There is global consensus that universal health coverage (UHC) will ensure that citizens are provided with access to adequate and appropriate healthcare (WHO, 2010; Evans, Marten and Etienne, 2012). Health insurance is an important potential tool in achieving UHC and contributory health insurance can assist to extend potentially limited government finance for health (WHO, 2010). Although evidence on the impact of health insurance on health outcomes remains elusive, there exists increasing evidence that health insurance helps households mitigate the effects of catastrophic health expenditure, while also improving access to healthcare (Finkelstein et al., 2012; Baicker et al., 2013).

In South Africa medical schemes are the main vehicle¹² through which the formal risk pooling of health expenditure occurs (Ramjee et al., 2014). It takes the form of private, not-for-profit health insurance, often employment based, of which the risk pooling function is enhanced by three legislatively protected principles: community rating,¹³ open enrolment and the provision of a defined package of minimum benefits (McLeod and Ramjee, 2007). By the end of 2013, 8.8 million South Africans (Council for Medical Schemes, 2014) or approximately only 17%¹⁴ of the population were medical scheme members. Apart from a small additional group using mainly private outpatient health services, the remainder of the population are reliant on the public health system.

Household survey data on health seeking behaviour and healthcare access demonstrate that medical schemes have played an “important mediating role...in accessing higher-quality private care” (McLaren, Ardington and Leibbrandt, 2013: 11). As Van den Heever (2012: S5) argues, “private [financing] systems play an important role in deepening coverage by mobilising revenue from income earners for health services over-and-above the horizontal extension role of public systems and related subsidies”. Van der Berg and McLeod (2009) estimated that between R176 billion and

¹² Hospital cash plan insurance is the only other form of insurance that provides insurance cover for medical expenses. The benefits of these products are generally limited to cover for hospitalisation. While households are able to purchase hospital cash plan insurance, this insurance does not indemnify (provide cover at actual cost) their medical expenses and there is currently much uncertainty about the legality of these products.

¹³ As opposed to risk rating. The price of medical scheme membership is not dependent on the risk profile of the member.

¹⁴ Council for Medical Schemes total calculated as percentage of estimated total South African population of 52.98 million as per Statistics South Africa Mid-Year Population Estimates (2013).

R251 billion was needed to transition South Africa into UHC through a National Health Insurance scheme (NHI)¹⁵ as then envisaged in an African National Congress (ANC) policy document. At a public health expenditure level of only R84 billion in 2009, this implied a significant funding shortfall which would necessitate large payroll taxes for the employed or large increases in general taxation levels (Van der Berg and McLeod, 2009). Given the potential funding shortfall, medical schemes and consumers' willingness to contribute to medical schemes could form an integral component of South Africa's move towards universal health coverage.

South Africa has not only one of the most unequal income distributions (Leibbrandt et al., 2010), but also has high levels of inequality in health outcomes between race groups. The latest racial breakdown of infant mortality rates is for 2002 and it shows a marked disparity: the infant mortality rate was 7 per 1,000 live births for white South Africans, while for the black population the rate was much higher at 67 per 1,000 live births (Coovadia et al, 2009). Inequality in health outcomes is also visible in the extreme variation of maternal mortality between provinces (Moodley, 2014). Coovadia et al. (2009: 817) notes the "almost three-fold difference in infant mortality between middle-class areas and squatter settlements" in the single geographic location of metropolitan Cape Town as identified in the research of Groenewald et al. (2008). Black South Africans are more than three times as likely as white South Africans to live more than 5kms away from a primary healthcare facility and significantly less likely to report having had a recent health consultation (McLaren, Ardington and Leibbrandt, 2014).

Medical schemes' role in providing access to healthcare becomes even more pertinent in the context a country which is not only characterised by health inequalities, but also has a public healthcare system which is under-funded compared to the private sector (McIntyre et al., 2007) and understaffed in terms of certain categories of providers such as medical specialists (Coovadia et al., 2009; Econex, 2010a). This results in rationing in the form of long waiting times at public healthcare facilities as reported in household surveys and waiting lists ("queuing") for medical procedures (Econex, 2010b).

Private health insurance can assist in decreasing pressure on public health systems, particularly for elective care (Tapay and Colombo, 2004). The enhanced provider choice and competition which are likely to result from greater access to private health financing may over the longer term also improve

¹⁵ Although a policy paper on the government plans for NHI was released for public discussion in 2011 (Department of Health, 2011a), no official follow-up proposals have since been made public. A review of the first eighteen months of efforts to test components of the NHI proposals indicates health system reforms under the NHI banner have until now mainly focused on the delivery of health services rather than experimentation with financing mechanisms (Matsotso and Fryatt, 2013).

the functioning of South Africa's polarised private and public healthcare markets. It is thus important to explore the degree to which the private health insurance market can be expanded to lighten the burden on the public system in at least the interim phase before full implementation of NHI.

The chapter starts by considering shifts in racial composition of the medical schemes market during the period 2000-2011 and I demonstrate that the market size and improvements in racial equity of access to medical schemes is limited by labour market dynamics and regulatory constraints. In order to identify the limits of the market, the correlates of *current* medical scheme membership are identified, such as access to a formal labour market job, affordability, health risk and perceptions of the quality of care. Given the limited recent growth in the medical scheme market, the potential for further growth in this market is explored through various affordability and regulatory relaxation scenarios.

2.2 The post-apartheid medical schemes landscape

In this section I provide an overview of some of the socio-demographic characteristics of the post-apartheid medical schemes landscape. Survey data was used for the analysis, because it contains characteristics of members that are not available in administrative data. This allows for consideration of how the racial composition of medical schemes has changed over time and the role of employment status in medical scheme membership. The use of race as a category to consider differences in medical scheme membership here is not intended to reinforce differences between South African sub-groups, but rather as a tool to monitor progress in the reduction of structural inequity given the continuing high correlation between race, class and income in South Africa. The use of employment status in the analysis is important as it provides some indication of the degree to which a reduction in inequality in access to healthcare is occurring through labour market shifts. The key regulatory barriers constraining the growth of the medical schemes market are also briefly mentioned.

Tables 2.A1-2.A6 (in Appendix A to this chapter) provides a summary of key data on racial and labour market shifts. Some of these tables are referred to in the discussion below.

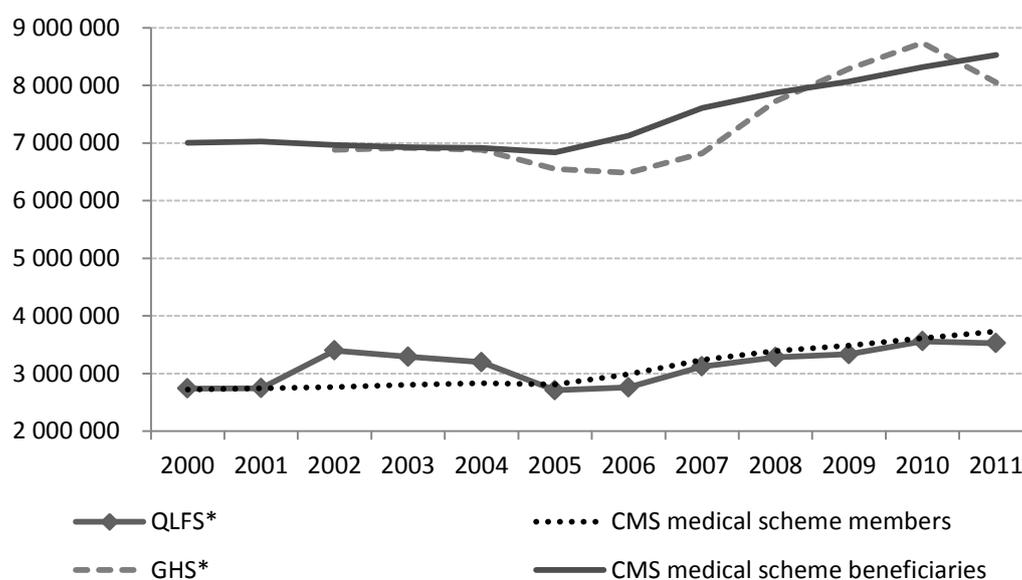
Figure 2.1, below, summarises the parallels between medical scheme cover data from labour force surveys and administrative data, on the one hand, and a household survey and administrative data, on the other hand. Data from a combination of the Labour Force Survey (LFS) (2000-2007) and Quarterly Labour Force Survey (QLFS) (2008-2011) mirror administrative data from the Council for Medical Schemes (CMS) on the total number of medical scheme members, while data from the General Household Survey (GHS) is closely aligned to administrative data on the total number of

medical scheme beneficiaries (members and their dependents)¹⁶. In some years survey data provide estimates that exceed the administrative data, while in other years the survey data provides under-estimation. In 2002 to 2004 the deviation of the LFS from administrative data was very large, with the survey over-estimating medical scheme membership by about 630,000 in 2002, almost 500,000 in 2003 and 360,000 in 2004. By 2005 the gap had reduced to 102,000 with the LFS in this instance providing an under-estimate of medical scheme membership.

Figure 2.1: Comparing trends in medical scheme membership across CMS and survey data

Source: The combined Labour Force Surveys (LFS) and Quarterly Labour Force Surveys (QLFS), various Council for Medical Schemes (CMS) annual reports (2001-2012) and the General Household Surveys (GHS)

**The Labour Force survey is a bi-annual survey while the Quarterly Labour Force Survey is conducted quarterly. To provide annually comparable data, data collected in September for both the Labour Force Survey and Quarterly Labour Force Survey were used.*



The comparison shows that the surveys do not perform poorly at capturing trends in main member and total medical scheme membership between 2000 and 2011. There is, for instance, no evidence that the surveys systematically under-capture a large proportion of members. This provides additional confidence in the use of these surveys for this purpose.

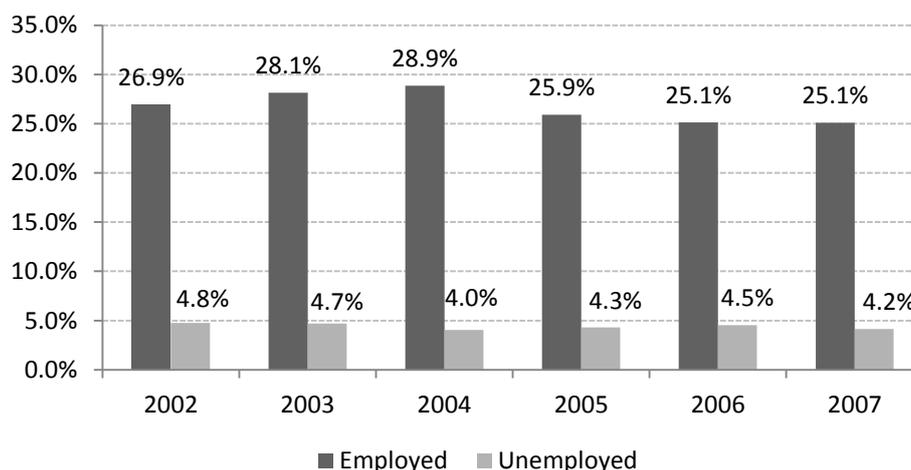
¹⁶ Where there are divergences between the QLFS and CMS data (on the one hand) and the GHS and CMS data (on the other hand), this is likely to be due to two possible reasons: the timing of the surveys relative to when administrative data is typically reported for (membership as on 31 December of any year) and the weighting of surveys. All in all, however, the survey data still approximates the administrative data relatively well and are therefore used in the analysis.

Various data sources confirm an increase in the total medical scheme membership, albeit small, from 2006 (Figure 2.1). The modest increase in both total medical scheme membership and membership of, in particular, black South Africans has been attributed mainly to the launch of the Government Employees Medical Scheme (GEMS) in 2006 and the extension of cover from this scheme to previously uncovered medical scheme members (Van Eeden, 2009).

Figure 2.2, below, shows the large differences in the prevalence of medical scheme cover between the employed and unemployed in South Africa. In 2007, the employed were six times more likely than the unemployed to have medical scheme cover¹⁷. While medical scheme cover can be purchased by anyone, employment forms the gateway to medical scheme cover for most.

Figure 2.2: Medical scheme cover amongst the employed vs. unemployed

Source: General Household Surveys, 2002-2007



Total membership by race provides information about the degree to which South Africans formerly excluded from many opportunities are gaining access to medical scheme cover (Table 2.A2). The provision of medical scheme cover by employers to black employees grew by 82% from only 1.1 million to almost 2.0 million during the 2000-2011 period (Table 2.A3). By 2010, according to the GHS, black South Africans represented 46.3% of total medical scheme membership. Despite this increase in total membership of black South Africans through mainly employment-funded coverage, a far lower percentage of black South Africans relative to white, Indian and coloured South Africans are members of medical schemes. Data from the GHS show that, in 2002, only 8.0% of black South

¹⁷ The nature of the labour market data collected in the GHS does not allow for the repeat of this type of estimate for more defined categories of employment, e.g. skilled and unskilled employment. While the LFS captures detailed employment data it does not collect data on medical scheme membership for the unemployed, as the medical scheme membership question is asked in the context of employment benefits.

Africans were members of medical schemes. The data appears to show an upward trend, though the survey nature of the data contributes to some fluctuations which make it difficult to distinguish trend from measurement error. In 2010, membership amongst blacks was measured at 10.2%, but the measure again declined to 8.8% in 2011. Much can therefore still be done to expand coverage to previously excluded population groups.

The change in medical scheme cover for black employees can be considered the result of two dynamics: more black South Africans moving into formal positions that are typically associated with medical scheme cover in both the public and private sectors (Table 2.A3) and the extension of medical scheme cover as an employment benefit to previously uncovered black South Africans working in the public sector.

In contrast, medical scheme cover for white South Africans obtained through employers decreased by 23.2% between 2000 and 2011, moving from cover for almost 1.2 million employees to cover for only 915,000. Much of this decrease can be attributed to a decrease in employment of whites in the public sector¹⁸ (see Table 2.A4).

Using data from the CMS, McLeod (2012b) estimates that 36% of all medical scheme members in 2009 were covered by some public sector-related medical scheme. The portion of public sector employees with medical scheme cover has gradually been increasing. Between 2007 and 2008, total public sector employees who indicated that their employer provides medical scheme cover increased from 59% to 73%, stabilising around 77% in 2011. At the same time, medical scheme cover for black formal sector employees also increased from 57% in 2007 to 70% in 2008. By 2011, 73% of black public sector workers indicated that they had some form of medical scheme cover through their employer (Figure 2.3).

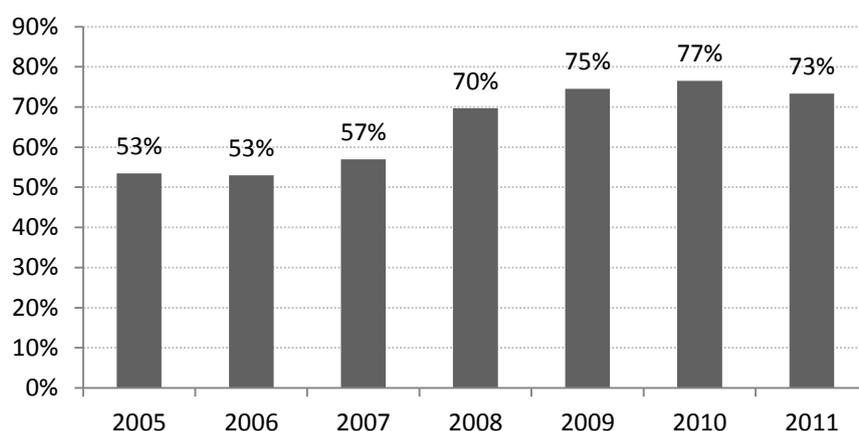
The public sector's role in the provision of medical scheme cover, particularly through the establishment and extension of GEMS, has been a major driver of employment-based medical scheme cover increases. Take-up of cover through GEMS has, however, been lower than anticipated. A study on a sample of government employees in the health and education sectors found that, despite the government medical scheme requiring no contribution from the lowest salary level of government employees, more than half of this group were uninsured by the survey year (Govender et al., 2013). The percentage of employees that chose to remain uninsured was highest amongst male, black and coloured employees and amongst employees with lower education and income levels. Government employees were seemingly discouraged from enrolling in the government

¹⁸ Employers in the private sector seemed to have decreased cover provided to employees and it negatively affected coverage of mainly white employees.

scheme by factors such as affordability, perceived administrative complexity of registration and even lack of access to information about the differentiated scheme benefit options (Govender et al, 2013).

Figure 2.3: Percentage of black employees with medical scheme cover in the public sector

Source: Labour Force Survey and Quarterly Labour Force Survey, September 2005-2011



Data on total medical scheme membership (i.e. including family members or dependents in addition to the main member) from the GHS show that white medical scheme membership stayed almost constant between 2000 and 2011, despite a decrease in cover provided by employers (Tables 2.A1 and 2.A4). In contrast, total black membership grew less than expected, despite the fact that medical scheme cover from employers almost doubled.

Growth constraining regulatory barriers:

The majority of medical scheme cover originates from employers as an employment benefit. In the past this cover was typically extended through “closed”, employer-based funds only open to employees and their family members. Only since the early 1980’s did “open” funds start to play a more important role in medical scheme coverage (Söderlund and Hansl, 2000). There are currently 87 medical scheme funds in South Africa, 24 of which are open schemes (Council for Medical Schemes, 2014). Changes in the size and socio-demographic composition of membership of medical schemes not only have to be understood relative to labour market and other socio-demographic changes in the country, but also have to be interpreted relative to regulatory and structural changes in the market itself¹⁹.

¹⁹ For a discussion of the history of regulation of medical schemes in South Africa, see Appendix B to this chapter.

In recent years, while awaiting more detailed regulatory proposals on the possible nature of a NHI scheme, there have been increasing concerns about the sustainability of the medical schemes market. Ramjee et al. (2014: 96) describes these concerns as follows:

“There is a view among stakeholders that there has not been sufficient regulatory attention paid to the current stability, sustainability and affordability of medical schemes. As such, these stakeholders feel that not enough is being done to ensure that the past and present criticisms of the medical scheme environment...continue”.

Two aspects are central to these concerns: firstly, the imperfect (and therefore arguably unsustainable) implementation of the guiding principle of social solidarity during the medical schemes re-regulation phase²⁰; and, secondly, the cost implications of the implementation of mandatory or Prescribed Minimum Benefits (PMBs). A key component of the imperfect implementation of social solidarity was failure to implement the envisaged risk equalisation fund (REF) to stabilise the industry (Ramjee et al., 2014). More details on both these concerns are provided in Appendix B to this chapter.

The cost implications of PMBs are exacerbated in an environment characterised by the absence of gatekeeping by general practitioners and a strong referral system (ERSA, 2014), while the absence of industry-wide tariffs or price guidelines for healthcare providers has also been noted as a concern (Ramjee et al., 2014).

Regulatory barriers help set the limits of the current market in terms of the affordability of medical schemes. The roles of affordability (or relative income) and other correlates of cover in the current market are explored in the next two sections.

2.3 The determinants of health insurance demand²¹

The demand for health insurance is influenced by both the demand for healthcare and the demand for financial risk management tools to assist in meeting the costs of ill health. Arrow (1963) distinguishes between two fundamental risks associated with health: the risk of ill health which implies potentially unexpected healthcare expenditure, decreased quality of life, loss of income

²⁰ See Appendix B to this chapter for a description of the various phases of the regulation of the medical schemes market in South Africa.

²¹ This section is informed by a literature search which used a variety of electronic search engines, including Google, Google Scholar, Econlit and ScienceDirect. Furthermore, some of the most prominent journals in health economics (Journal of Health Economics and Health Economics) were also searched directly. The main search terms that were used include “determinants” or “correlates” in combination with the phrases “health insurance” or “voluntary health insurance”. Once initial important sources were found, their bibliographies were also consulted for additional sources.

generating ability and potentially even death; and complete, partial or delayed recovery. The financial losses associated with the former risk can be partly managed through health insurance and loss of income insurance, while the latter risk can be partly managed with disability insurance. When considering the determinants of health insurance, it is thus necessary to identify the factors that influence both the demand for healthcare and the demand for a health expenditure risk management tool. Here I briefly consider the available South African and international literature on three possible demand determinant categories: employment and income; risk and health status; and the perceived quality of health services. Other factors which have been considered in the literature are also highlighted.

2.3.1 Employment and income

Employment status influences both access to social security benefits and access to income. A positive association between higher income and health insurance coverage has been confirmed by multiple studies of the United Kingdom's voluntary private health insurance market (Besley, Hall and Preston, 1999, King and Mossialos, 2005; Propper, 1993; Propper, Rees and Green, 2001). In the United Kingdom, voluntary private health insurance market membership has been found to be positively associated with paid employment and professional or managerial employment (King and Mossialos, 2005). Propper (1993) finds self-employment to be negatively associated with captive preferences in favour of the public health system in the United Kingdom, i.e. the self-employed are less likely to be captive to the public sector health system in their choice set and would therefore be open to purchasing private health insurance. In explaining this association, Propper (1993: 299-300) argues that the self-employed are probably "more likely to consider insurance as a way of avoiding the income loss associated with being on a waiting list for medical care".

Household income and the affordability of medical scheme cover are critical determinants of medical scheme membership. Affordability is influenced by the cost of medical scheme cover, household income and household size.

In the South African context, it is likely that higher income allows individuals to opt out of the public system and that the decision to opt into medical scheme membership is therefore a secondary effect of higher income. Grobler and Stuart (2007) focus on the factors associated with selection into specific categories of healthcare (public, private or no care) in South Africa. They find income to have a significant influence on the category of healthcare selected. Individuals from all income quintiles use private healthcare, but individuals up to the fourth quintile are found to be more likely to use public healthcare, while individuals in the fifth quintile mostly use private healthcare. The demand

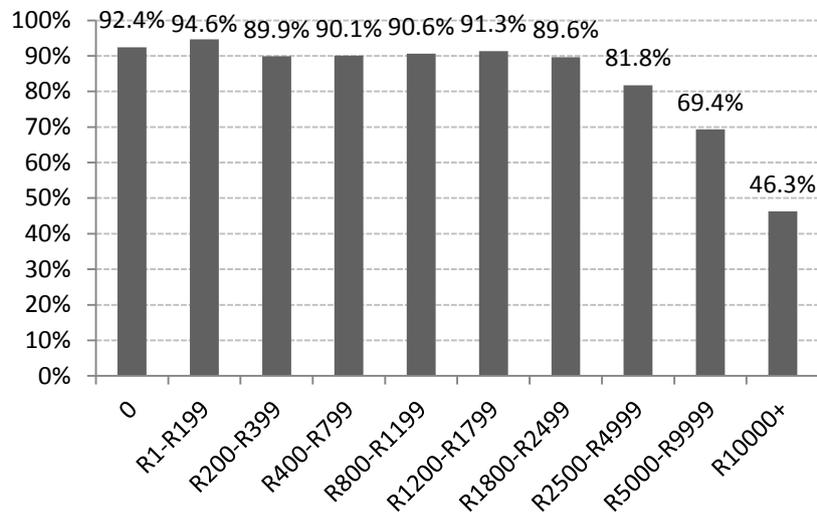
for health insurance is derived from the demand for healthcare. Although the study by Grobler and Stuart (2007) does not focus on medical scheme membership, the same association between income and private healthcare is like to hold for medical scheme membership.

Medical scheme membership in South Africa provides access to mainly private healthcare and one would therefore expect a similar relationship as between income and healthcare provider choice. Kirigia et al. (2005) use a binary logit model to consider the relationship between health insurance cover and demographic, economic and education variables, as obtained from the 1994 Health Inequalities Survey. Their analysis focuses on the factors that influence health insurance coverage amongst South African women as the survey mainly sampled women. Health insurance cover amongst South African women was found to be positively and significantly ($P < 0.05$) associated with higher income, being employed and being employed in a white collar occupation (Kirigia et al., 2005).

In 2009, the GHS included a special section with questions on why households do not have medical scheme cover, if they indicated that none of the household members were covered by a medical scheme. Large percentages of households cited affordability reasons as the main constraint to medical scheme access. Figure 2.4, below, shows the proportion of households with no medical scheme cover who report affordability as an access barrier by income category - only 46.3% of households cite affordability reasons as the main medical scheme access constraint in the group of households with monthly expenditure of R10,000 or more, compared to 89.6% of households with monthly expenditure of R1,800-R2,499.

Figure 2.4: Percentage of households that have no medical scheme membership due to affordability reasons by household expenditure category

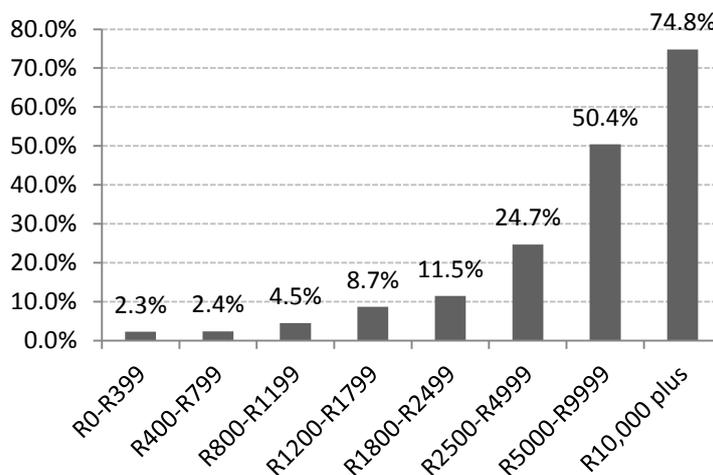
Source: General Household Survey 2009



In the absence of reliable expenditure data which allows for calculation of an affordability measure, variation in medical scheme cover by household expenditure provides some indication of affordability. Figure 2.5 indicates that there is a clear increase by household expenditure category in the percentage of total medical scheme members that fall within a specific category. More than 56% of medical scheme members indicated that their monthly household expenditure totalled R5,000 or more.

Figure 2.5: Percentage of total medical scheme members by household expenditure category

Source: General Household Survey, 2009



2.3.2 Risk and health status

Information asymmetries in insurance markets could lead to both adverse selection (Akerlof, 1970) and moral hazard as identified by Arrow (1963) and Pauly (1968). While Pauly's (1968) theoretical framework for insurance implies a net welfare loss due to moral hazard²², in Nyman's (1999) framework in which insurance purchases are responsive to income²³, insurance becomes a welfare transfer from the healthy (and potentially wealthy) to the sick. Nyman's (1999) view of insurance as a redistributive mechanism, however, is reliant on a large or universally insured market and may have less application value to South Africa where the poor are not active participants in the medical schemes market.

Recent studies, however, have found that risk selection may also occur in the opposite direction, with persons offering good or positive risks advantageously selecting into insurance markets (Finkelstein and McGarry, 2006; Fang, Kean and Silverman, 2008). The nature of risk selection (adverse or advantageous) is likely to vary with the type of insurance market (Finkelstein and McGarry, 2006), with evidence emerging that positive or advantageous selection is also occurring in health insurance markets (Fang, Keane and Silverman, 2008).

The decision to purchase health insurance and then use it in accessing health services is not independent from health status. Health status is, however, only one dimension of risk in health insurance markets. Other dimensions include risk preferences or private knowledge (not accessible to the insurer) on the likelihood of risky behaviour once insured (Fang, Keane and Silverman, 2006). Collectively, these dimensions influence the likely health expenditure which the insurer will have to incur on the insured.

While individuals with poor health may be more likely than their healthier counterparts to purchase health insurance, it is also possible that the insured may simply report poorer health status compared to the uninsured because they have better access to healthcare. Insurance provides free-of-charge or, at the least, discounted access to a variety of healthcare providers. This may lead to greater utilisation of healthcare or the use of different healthcare providers compared to the uninsured which may, in turn, lead to better diagnosis of existing underlying healthcare conditions. Breyer, Bundorf and Pauly (2011:746) describe this endogeneity dilemma when using health status as a proxy for risk from the insurer's perspective as follows:

²² Pauly's (1968) framework did not account for the influence of income on likelihood of insurance purchase.

²³ The likelihood of insurance purchases increases with income.

“The potential endogeneity of measures of health represents an additional concern. To the extent that insurance coverage increases the rate of diagnosis of conditions (in the case of indicators of chronic conditions) or the effectiveness of treatment (in the case of self-reported health status), these studies will overestimate the extent of risk-based selection.”

Contradicting evidence on the nature of the association between health status and insurance coverage makes a final conclusion on the nature of this relationship difficult. Some studies find no clear association (King and Mossialos, 2005), while a recent review of available literature on the demand determinants for private health insurance in universal health care systems concluded that, in most studies, the insured tend to be in equivalent or better health states than their uninsured counterparts (Kill, 2012). However, when the health expenditure of individuals predicted by the insurer, conditioned on health status, is used as health risk measure, at least two studies find a positive association between higher expenditure (implied poorer health risk) and insurance coverage in, respectively, Australia and the United States (Fang, Keane and Silverman; Bundorf, Herring and Pauly, 2010). It is increasingly recognised that the nature of the health status measure may have an influence on the sign of the association (Breyer, Bundorf and Pauly, 2011). Thus, for example, in a group of families who declined take-up of health insurance offered by employers in the United States, individuals with poor self-assessed health were found to be more likely to decline cover, while those with fewer diagnosed chronic conditions were also more likely to decline cover (Benard and Selden, 2006).

Given the potentially complex issue of endogeneity in the relationship between reported health status and insurance usage, it is difficult to establish the true direction of causality if healthcare utilisation is not taken into account and without the use of causal methods or experimental data.

In South Africa, medical scheme members report a significantly ($P < 0.01$) higher prevalence of chronic diseases and also more frequently report being ill or injured in the four weeks preceding the survey than their uninsured counterparts (Table 2.A5). For instance, in 2009 17.1% of medical scheme members reported having at least one of a number of chronic diseases compared to only 12.3% of the uninsured, while 23% of medical scheme members said they were ill or injured in the six weeks preceding the survey compared to only 17.6% of the uninsured.

It is not clear whether this is due to risk selection or whether medical scheme members simply have better access to health services, allowing these conditions to be diagnosed. Experimental data or causal methods are required to explore the direction of causality.

2.3.3 Consumer choice and healthcare acceptability

Voluntary health insurance is able to provide the insured with access to private health services. This includes services not typically provided by the public sector or at potentially higher levels of user acceptability. It is important to emphasise that user acceptability in the context of health services is a broad concept consisting of many dimensions, which may range from user experiences of the diagnosis and interaction with health staff, the standard of healthcare facilities (cleanliness, maintenance, etc.), waiting times (some may refer to this as convenience) and even health staff attitudes. User acceptability differentials between the services provided by the public and private sector can potentially form an important explanation for the demand for health insurance.

The international literature on correlates of voluntary private health insurance membership uses the term “quality” rather than user acceptability to describe users’ experiences and perceptions of healthcare. Waiting lists or waiting times for accessing health services are often used interchangeably as an easily measurable health service quality proxy. Both indicators can be considered signals of rationing in a public health system.

In the absence of data on waiting times for healthcare, data on waiting lists for certain procedures are used and emerge as a positive and significant determinant for private health insurance membership in the United Kingdom (Besley, Hall and Preston, 1999). However, once data on both waiting lists and waiting times are considered, waiting lists become an insignificant variable with only waiting times having a significant impact on the demand for health insurance (Johar et al., 2013). King and Mossialos (2005) also find longer waiting times for healthcare to be positively associated with private health insurance coverage in the United Kingdom.

An alternative approach is to use the healthcare acceptability gap or the difference in perceptions on various dimensions of healthcare acceptability (or quality) for private services compared to public services. In this case, a positive value indicates higher or better quality perceptions of the private health sector relative to the public sector. A quality gap variable constructed using the aforementioned approach was found to be positively and significantly associated with private health insurance coverage in Catalonia, Spain (Costa and Garcia, 2003).

2.4 Correlates of medical scheme demand

2.4.1 Data

South Africa has limited administrative data on the characteristics of medical scheme members. While publically available administrative data from the CMS allows for the tracking of trends in total membership and beneficiary numbers, as well as average age and distribution across age bands, no

further information on medical scheme members is made available by the sector regulator. This study therefore uses two nationally representative datasets to examine the correlates of medical scheme demand: data from the LFS (2000-2007) and the QLFS (2008-2011) pooled into one dataset and pooled data from consecutive GHSs (2002-2011). These two pooled datasets allow for, respectively, robust estimation of the correlates of having medical scheme cover through an employer when employed for those of working age (LFS/QLFS) and the correlates of total medical scheme membership, which includes main members and their dependents (GHS).

The two sets of labour force surveys ask all respondents who indicate that they are employed whether their employer provides medical scheme cover. The LFS/QLFS thus allow for the teasing out of the variables associated with having medical scheme cover for all *employed* South Africans, as both the LFS and QLFS are nationally representative surveys of working-age individuals. In contrast, the GHS asks all respondents, irrespective of age, whether they have medical scheme cover. As the GHS is also a nationally representative survey, the population of interest thus is *all South Africans*.

The pooled datasets were created by merging data for different years for those variables that had been most consistently recorded across years, as not all questions were asked every year. There were therefore a number of variables that were not available for the full period and when these variables are included in the analysis, much smaller samples were available for analysis.

The demand for health insurance or, more specifically, medical scheme cover, is likely to be influenced by factors that determine the demand for healthcare and those that are unique to the demand for health insurance. Below a brief explanation is provided of some of the variables used in the analysis and the likely direction of the relationship between these variables and the probability of being a medical scheme member. The selection of variables was informed by intuition as well as the international literature reviewed in Section 2.3.

Labour market variables:

The LFS/QLFS dataset provides a rich source of variables on the characteristics of the employed and the organisations by which they are employed. In the analysis I therefore included binary variables on formal employment, skilled employment, public sector employment and union membership. It is anticipated that being employed in the formal sector (relative to the informal sector) is likely to increase the probability of having medical scheme cover through the employer, while being in a skilled position is also likely to have a positive relationship with the probability of having medical scheme cover.

Socio-economic variables:

I used different variables to control for the socio-economic correlates of medical scheme demand. While the analysis did not control for any income variables in LPMs run with the QLFS/LFS dataset as this data had not been collected, per capita expenditure was calculated from a household expenditure variable in the GHS dataset to control for income and the concept of affordability.

In addition to the per capita household expenditure variable, I included three variables to control for access to services which are likely to capture aspects of socio-economic vulnerability not fully captured by the expenditure variable. These variables include dummies to control for residing in a formal housing structure, reporting access to clean water and having a flush toilet (as opposed to bucket or other toilet options). Due to the small number of indicators that are available across all years of the pooled dataset, it was not possible to control for the ownership of physical assets.

The number of children in the household is often included as a control variable in studies which consider the demand determinants for private health insurance. After considering eleven empirical studies which include children as a demand determinant, Kill (2012: 106) concludes that “there is no clear-cut association between the presence or number of children in the household and VPHI [voluntary private health insurance] status across countries or insurance types”.

Apart from implicitly controlling for household size through use of the per capita expenditure variable in the GHS dataset, I also included it as a separate variable in the LPMs for both the GHS and LFS/QLFS data, as household size may have an additional effect beyond only affordability concerns in decisions about medical scheme coverage.

In South Africa a minimum education level of matric (completion of high school), living in an urban formal area, a good perception of the surrounding residential environment and smoking have been found to be positively associated with health insurance cover (Kirigia et al., 2005). Household size, the use of contraceptives and consumption of alcohol, potentially both indicators of risky behaviour, were negatively associated with health insurance cover (Kirigia et al., 2005).

Most international studies find a positive association between level of education and private voluntary health insurance cover (Besley, Hall and Preston, 1999; Hall and Preston, 1989; King and Mossialos, 2005; Kill, 2012). I controlled for education by including binary variables for 12-14 years of education and 15 or more years of education in the LPMs using the LFS/QLFS data. For the LPMs using the GHS data the presence of someone in the household with an education of 15 or more years of education was controlled for. While education can also be included as a continuous variable, binary variables lend themselves to easier interpretation in LPMs.

Studies which consider the demand correlates for private health insurance often control for health behaviours which may be detrimental to health status such as smoking or alcohol consumption. While some studies find smoking to be positively associated with private voluntary health insurance (King and Mossialos, 2005), other studies (Propper, 1989) find no association. No such variables were included in my analysis as they were not available in the two datasets.

Demographic variables:

I controlled for the enduring effect of race with a dummy for white individuals. Aligned with the general medical scheme member structure discussed in earlier sections, I hypothesised the probability of having medical scheme cover remains linked to race. A male dummy was also included to control for gender.

According to Grossman (1972) the demand for healthcare is likely to increase with age since the stock of health depletes with ageing, implying an expected positive association between the demand for healthcare, insurance and age. It is possible, however, that individuals may want to hold insurance for certain health events for which they prefer to use private health services. In South Africa there exists anecdotal evidence of households or individuals obtaining medical scheme membership in anticipation of child birth and the medical services both the mother and child will consequently require (Econex, 2012).

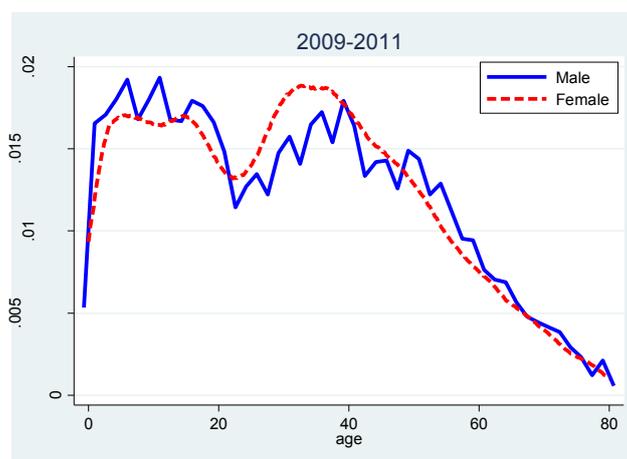
In international studies the relationship between age and health insurance cover is typically found to be positive (Propper, 1993; Propper, 2001). Some studies find that age is positively associated with voluntary private health insurance up to a certain threshold, and thereafter the association becomes negative or insignificant (Besley, Hall and Preston, 1999; King and Mossialos, 2005; Kill, 2012). In the context of these studies, age has a positive and statistically significant relationship with private insurance cover, largely as proxy for illness and awareness of risks rather than due to affordability considerations.

If scheme membership is determined mainly by need rather than affordability, one would assume membership to increase with age, but there may also be increased demand during the childbearing years of women. The age distribution of medical scheme beneficiaries in the GHS dataset indicates a distribution slightly contrary to this expectation. While there is indeed a sharp increase in medical scheme membership that coincides with the typical childbearing age period of women, this age period is also closely associated with the economically active years of both male and female members. Contrary to the expectation created by Grossman's conclusion on the direction of the age relationship, the age distribution displays a "twin peak" profile with the first peak occurring during

the childhood years but decreasing sharply around 18 years of age, followed by the start of the second peak around the start of the economically active years that in South Africa also coincide with the typical childbearing years of women (Figure 2.6, below)²⁴. Contrary to Grossman's (1972) view on the demand for healthcare increasing with the stock of age, medical scheme membership starts to rapidly decrease from around 50 years of age.

Figure 2.6: Kernel density function of age distribution of all medical scheme members: 2009-2011

Source: General Household Surveys



If the age distribution for the two largest race groups of medical scheme members, black and white members, is considered separately, two different pictures emerge. There is no such prominent twin peak in the age distribution visible for white members (Figure 2.7) and far more white older persons are members of medical schemes compared to black older persons (Figure 2.8), although medical scheme membership for white individuals also decreases with age²⁵. The twin peak distribution evident in the age distribution of all medical scheme members is even more prominent for black members (Figure 2.8) with a peak in membership for both male and female members shortly after 40 years of age followed by a sharp decline in membership. These differential patterns in the age distribution of white and black members are most likely driven by the greater prevalence of selective coverage within households amongst black households compared to white households.

²⁴ Similar kernel density curves for the other time periods of the pooled dataset can be found in Figures 2.A1 in Appendix A to the chapter.

²⁵ Similar kernel density functions for the other time periods in the pooled dataset can be found in Figures 2.A2 and 2.A3 in Appendix A to the chapter.

Figure 2.7: Kernel density function of age distribution of white medical scheme members for 2009-2011

Source: General Household Surveys

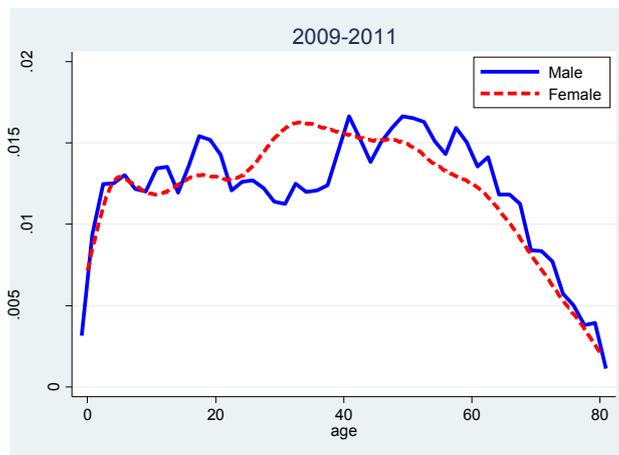
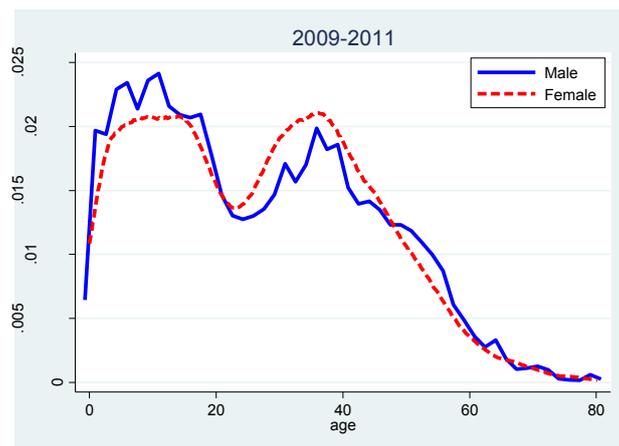


Figure 2.8: Kernel density function of age distribution of black medical scheme members for 2009-2011

Source: General Household Surveys



In the regression analysis, I controlled for age using binary variables for different age groups in order to capture possible threshold effects associated with certain age groups. This was done for both the QLFS/LFS and GHS data.

Risk, consumer choice and healthcare acceptability variables:

The GHS dataset allows me to control for health status using two health or health-risk related variables: reporting an injury or illness in the four weeks preceding the survey and the prevalence of chronic disease. The latter variable is constructed to identify individuals that indicated they have been diagnosed with one of a number of chronic diseases, including diabetes, hypertension, arthritis, asthma, HIV, cancer and an “other” category of chronic diseases. While these variables are potentially important in explaining the take-up of medical scheme cover, they are also prone to the

problem of endogeneity as explained in the section dealing with health risk as determinant of health insurance membership. Illness and chronic disease may have a higher likelihood of diagnosis amongst medical scheme members because of access to health insurance. No causal impact of these variables on the demand for medical scheme cover can therefore be inferred.

There is no variable in the dataset that directly measures acceptability of public healthcare services received. There is also no variable that allows for the easy merging in of administrative data on acceptability of public health services available in the area and match this with specific respondents on geographic level. I therefore used the responses to a question introduced in the 2009 GHS on why the closest health facility was bypassed to construct a binary variable to capture acceptability of public healthcare services on primary sampling unit²⁶ (PSU) level. If any one of a number of potential responses to the question on why the nearest healthcare facility was bypassed was provided, these responses were treated as indicative of low acceptability. The possible responses include: facilities not clean; long waiting times; opening times not convenient; drugs that were needed not available; staff rude or uncaring or turned patient away; and incorrect diagnosis. In the binary variable, these responses take on a value of zero. The number of responses indicating acceptable services (i.e. one values) is totalled on PSU level and converted to a proportion of total possible responses²⁷ on PSU level. If the proportion was equal to or exceeded 75% (alternatively, if 75% or more of respondents at any PSU level did not provide reasons relating to poor acceptability for bypassing the closest public healthcare facility), these PSUs were treated as if they provided healthcare of a sufficient acceptability level. As the PSU for each individual is known, I was able to control for acceptability of public healthcare services at the PSU level for individuals in the regressions through use of this binary variable.

I also controlled for an excessively long travel time (proxy for distance) to public healthcare facilities. This variable was constructed similar to the acceptability variable on PSU level. The responses of individuals who indicated they typically consult at a public healthcare facility and have to travel there for 30 minutes or longer were given a value of one. The responses of individuals who also consulted public healthcare facilities but typically travel for less than 30 minutes take a value of zero. The one value responses were aggregated at PSU level and converted to a proportion of total possible responses at PSU level. If the proportion of public healthcare users at PSU level who

²⁶ The lowest geographic level captured by the survey.

²⁷ Two versions of this variable were created. The first one incorporates responses (or lack thereof) on acceptability of healthcare services for all individuals across the income spectrum. In the second version responses of individuals in the first and fifth (bottom 20% and top 20%) household expenditure quintiles were excluded as these may suffer from health service quality perception and/or reporting bias issues. The results of both versions of these variables are reported.

reported travelling for longer than 30 minutes was equal to 75% or more (alternatively, if 75% or more of respondents who use public healthcare facilities indicated travel times of 30 minutes or more), these PSUs were treated as if having long travel times to public healthcare facilities and thus being less accessible public healthcare facilities. This allowed me to control for consumer choice in the context of decisions on which type of healthcare provider (public or private) to consult and therefore for the usefulness of medical scheme membership.

2.4.2 Methods

Wooldridge (2002) identifies linear probability models (LPMs) as a possible functional form for binary response specification. It is argued that these models “should be seen as a convenient approximation to the underlying response probability” (Wooldridge, 2002: 454).

I used LPMs to explore the relationship between being a member of a medical scheme and four categories of variables: labour market, socio-economic and demographic variables, as well as variables on health status and use of health services as described in Section 2.4.1 above. While both datasets contain most of the categories of required variables, the precise variables vary across the two datasets. The combined LFS/QLFS dataset does not contain any variables on health status, while the GHS dataset does. The LFS/QLFS dataset contains more detailed labour market variables (e.g. formal employment, skills level, public vs. private sector employment) while the GHS only contains a broad employment variable. LPMs were therefore specified for both datasets using as many as possible socio-demographic and health status, as well as health service variables.

LPMs are prone to the problem of heteroskedasticity, which is not the case with probit models. However, the potential of heteroskedasticity is likely to have a relatively small effect on the empirical results produced (Angrist and Pischke, 2008). Although the results are not reported here, as a robustness check I ran probit models in parallel to the LPMs using the same variables and the results did not differ in any meaningful way.

2.4.3 Results

The aim of the paper is to determine which generalisable characteristics are associated with medical scheme membership and what this implies for access to healthcare services. Due to endogeneity concerns relating to income and health status measures, coefficients cannot be interpreted as representing an estimate of the causal impact. Below, I discuss the variables estimated to be the strongest correlates of medical scheme cover, both for employees and all medical scheme members. The regression results for the LFS/QLFS and GHS datasets are reported separately and differences or similarities between these two sets of results highlighted where appropriate.

Factors associated with employer-based medical scheme cover (LFS/QLFS dataset):

Not all variables are available for every year. Consequently four different specifications were estimated – two estimates for each of two time periods which allows for observation of changes in coefficient size and signs when certain variables are added. The time period covers the years 2000-2007 as well as 2010-2011, years that contain variables related to labour union participation not covered in the others years; the second time period covers the years 2000-2005 and 2010-2011, which allows for inclusion of a variable controlling for rural/urban status.

Almost all variables included in the four LPM specifications of medical scheme cover through an employer are statistically significant (Table 2.1). In the second and fourth estimates a binary variable for union membership was added. The labour union variable is likely to suffer from endogeneity issues as it may be influenced by the fact that labour unions attempt to negotiate greater social security benefits for their members. I therefore proceed cautiously in interpreting the meaning of any association between labour union membership and medical scheme cover. In the last model a dummy variable for rural areas was added. The coefficient sizes of some of the largest correlates of medical scheme cover show large changes when the additional variables are added.

In the initial and third estimates, public sector employment and an education of 15 years or more (equivalent to having at least a degree qualification) have the largest association with medical scheme membership. These variables are followed in size by being employed by a firm with 50 staff or more, being a white employee and employment in a skilled position (as opposed to an unskilled or semi-skilled position). However, when a dummy for union membership is added in the second and fourth models, it emerges as the largest correlate of medical scheme cover. In fact, being a union member increases an employee's probability of having medical scheme cover by more than 25 percentage points. An education of 15 years or more increases an employee's probability of having medical scheme membership through their employer by 23.7 percentage points or more compared to not having completed matric.

Including the union dummy causes a number of shifts in other coefficients due to the web of relationships connecting the labour market variables. Adding the dummy increases the coefficient on membership for white workers further, while inclusion of the union dummy decreases the size of association of public sector employment with medical scheme cover. The rural variable added in the fourth model has a relatively small coefficient and in the direction expected, i.e. living in a rural area makes an employee less likely to have medical scheme cover. White employees have an increased likelihood of medical scheme cover of 16 to 19 percentage points.

Table 2.1: Output of linear probability model using LFS and QLFS data (2000-2011)

Variables	Dependent variable: Medical scheme cover			
	2000-2007, 2010-2011		2000-2005, 2010-2011	
	(1)	(2)	(3)	(4)
Skilled employment	0.136*** (0.00156)	0.132*** (0.00151)	0.139*** (0.00174)	0.132*** (0.00172)
Public sector employment	0.329*** (0.00162)	0.223*** (0.00166)	0.333*** (0.00178)	0.222*** (0.00188)
White	0.161*** (0.00162)	0.191*** (0.00157)	0.157*** (0.00181)	0.184*** (0.00180)
Education 12-14 years	0.140*** (0.00129)	0.123*** (0.00125)	0.147*** (0.00144)	0.122*** (0.00144)
Education 15 years plus	0.248*** (0.00253)	0.242*** (0.00244)	0.249*** (0.00281)	0.235*** (0.00278)
Male	0.0303*** (0.00108)	0.0189*** (0.00105)	0.0309*** (0.00121)	0.0193*** (0.00119)
Age 21-40	0.0306*** (0.00315)	0.0164*** (0.00304)	0.0341*** (0.00356)	0.0143*** (0.00348)
Age 41-60	0.0769*** (0.00325)	0.0497*** (0.00314)	0.0810*** (0.00367)	0.0475*** (0.00360)
Age 61 plus	0.0349*** (0.00468)	0.0246*** (0.00452)	0.0384*** (0.00531)	0.0225*** (0.00519)
Married	0.0629*** (0.00115)	0.0460*** (0.00112)	0.0666*** (0.00129)	0.0484*** (0.00127)
Household size	-0.00459*** (0.000347)	-0.00253*** (0.000335)	-0.00553*** (0.000384)	-0.00332*** (0.000378)
Children	0.00110* (0.000607)	1.36e-05 (0.000585)	0.00119* (0.000675)	0.00186*** (0.000666)
Firm size 21-49 employees	0.108*** (0.00153)	0.0698*** (0.00149)	0.105*** (0.00171)	0.0660*** (0.00171)
Firm size 50 plus employees	0.234*** (0.00125)	0.163*** (0.00126)	0.237*** (0.00138)	0.161*** (0.00143)
Union member		0.253*** (0.00136)		0.258*** (0.00155)
Rural				-0.0376*** (0.00137)
Year controls (2000 as reference year)	YES	YES	YES	YES
Constant	-0.0726*** (0.00367)	-0.0681*** (0.00354)	-0.0768*** (0.00405)	-0.0509*** (0.00402)
Observations	461,577	461,577	378,146	361,380
R-squared	0.366	0.410	0.370	0.414

Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

As the medical scheme cover question is only asked to employed individuals, the available age variable is limited to the economically active age group of 15-65. The three binary age variables show association patterns slightly contrary to expectations²⁸. The 15-20 age range serves as the reference category for comparing the coefficients of the other age categories. Many in this age group would not yet be economically active. Relative to the reference variable, being in the age group 41-60 years is the age category that matters most for having medical scheme cover. It is associated with an increase in the probability of medical scheme cover of between 5 and 8 percentage points, considerably larger associations than that found for the other age categories. I expected that the age group of 61 and older who are still employed would have a higher probability of medical scheme cover, given the close association between poor health and ageing. However, coverage amongst this older group seems to have not reached the levels that have become more common in recent years.

Across all specifications, being formally (as opposed to informally) employed has a relatively small association with having medical scheme cover.

Factors associated with cover for all medical scheme members (GHS dataset):

Similar to the analysis of the LFS/QLFS data, different regression specifications were run for different year groups due to not all variables being available for all years. The period 2009-2011 is both the most recent time period in the dataset and the most comprehensive in terms of variable availability. I therefore focus most of the discussion on these results. To compare results for this time period to the other years covered in the dataset, regressions were also run with fewer variables for 2002-2008 and 2002-2011 (the full time period). The results for this latter group of regressions can be found in Table 2.A6 of Appendix A. The greater availability of variables in the period 2009-2011 allowed me to control for urban rural/status, the reporting of any type of chronic disease (health status), perceived acceptability of health care services and distance to the health facility consulted if ill. It is not possible to include these variables for the other time periods. There is also data on PSU level available for this period, which allows for the construction of variables that control for the nature of public health services (acceptability and distance) at this level.

Almost all variables included in the different specifications of membership for all medical scheme members are statistically significant (Table 2.2, specifications 1-7). The only variables that are statistically insignificant is reporting having access to clean water for household use and reported

²⁸ I also ran versions of the same specifications where I controlled for age using dummies for age for individual years (the results are not reported here). This did not make a difference to the results obtained although it allowed for the detection of coefficient sign changes on a more granular level.

long distance (travel time of 30 minutes or more) to public healthcare facilities at PSU level. This is discussed later in the section.

The variables with the largest associations with medical scheme membership are the white race dummy and having a degree (15 years or more of education). Across the seven specifications for the 2009-2011 period (Table 2.2), being white (as opposed to black, coloured or Indian) is associated with an enhanced probability of medical scheme cover of 29.3-29.6 percentage points. Living in a household with someone who has a degree or three-year post-school qualification is associated with a probability of medical scheme cover of about 26 percentage points higher than for households without a degreed or similar member.

The age group of 61 and older is used a base category²⁹. The age categories of 0-4, 5-14 and 15-20 have the largest association with the dependent variable. Being a child (0-4, 5-14 or 15-20) is associated with a probability of medical scheme cover of at least 6.4 percentage points higher than the reference age group. Interestingly and contrary to the expectation of a higher probability of medical scheme cover in the childbearing years, falling in the 41-60 age category is associated with a medical scheme coverage probability of almost twice the size of falling in the 21-40 (childbearing) age category.

The log of per capita expenditure has a sizeable positive association with the probability of having a medical scheme. Two access to services variables (formal housing and flush toilet) are also positively and significantly associated with medical scheme cover. Reporting access to clean water for household purposes has a negative but insignificant association with medical scheme cover. This is most likely due to the large prevalence of access to clean water in South Africa.

Household size is statistically significant and positively associated with medical scheme cover. Regressions with the GHS data also control for household size through inclusion of per capita household expenditure by adjusting expenditure for the number of household members, irrespective of whether adults or children. An alternative approach would be to control for number of adult and child household members separately.

²⁹ I also ran versions of the same specifications where I controlled for age using dummies for age for individual years (the results are not reported here). This did not make a difference to the results obtained although it allowed for the detection of coefficient sign changes on a more granular level.

Table 2.2: Output of linear probability model using GHS (2009-2011)

Variables	Dependent variable: medical scheme cover						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log of per capita household expenditure	0.101*** (0.000647)	0.101*** (0.000648)	0.101*** (0.000648)	0.101*** (0.000650)	0.101*** (0.000650)	0.101*** (0.000650)	0.101*** (0.000650)
Formal housing	0.0397*** (0.00152)	0.0387*** (0.00153)	0.0386*** (0.00153)	0.0389*** (0.00154)	0.0390*** (0.00154)	0.0391*** (0.00154)	0.0391*** (0.00154)
Clean water access	-0.00418** (0.00207)	-0.00210 (0.00211)	-0.00240 (0.00211)	-0.00282 (0.00212)	-0.00254 (0.00212)	-0.00242 (0.00213)	-0.00222 (0.00213)
Flush toilet	0.0353*** (0.00140)	0.0416*** (0.00184)	0.0416*** (0.00184)	0.0421*** (0.00184)	0.0422*** (0.00184)	0.0423*** (0.00185)	0.0423*** (0.00185)
Male	-0.0122*** (0.00115)	-0.0122*** (0.00115)	-0.0119*** (0.00115)	-0.0115*** (0.00116)	-0.0115*** (0.00116)	-0.0115*** (0.00116)	-0.0115*** (0.00116)
Married	0.0602*** (0.00158)	0.0602*** (0.00158)	0.0602*** (0.00158)	0.0595*** (0.00159)	0.0595*** (0.00159)	0.0595*** (0.00159)	0.0593*** (0.00159)
Age 0-4	0.0794*** (0.00294)	0.0798*** (0.00294)	0.0811*** (0.00294)	0.0861*** (0.00311)	0.0862*** (0.00311)	0.0862*** (0.00311)	0.0860*** (0.00311)
Age 5-14	0.0824*** (0.00263)	0.0827*** (0.00263)	0.0849*** (0.00264)	0.0894*** (0.00281)	0.0895*** (0.00281)	0.0895*** (0.00281)	0.0892*** (0.00281)
Age 15-20	0.0636*** (0.00280)	0.0638*** (0.00280)	0.0664*** (0.00281)	0.0708*** (0.00297)	0.0707*** (0.00297)	0.0707*** (0.00297)	0.0704*** (0.00296)
Age 21-40	0.00348 (0.00248)	0.00384 (0.00248)	0.00601** (0.00249)	0.00989*** (0.00262)	0.01000*** (0.00262)	0.0100*** (0.00262)	0.00981*** (0.00262)
Age 41-60	0.0206*** (0.00266)	0.0208*** (0.00266)	0.0216*** (0.00266)	0.0234*** (0.00269)	0.0235*** (0.00269)	0.0235*** (0.00269)	0.0233*** (0.00269)
Household education 15 years plus	0.259***	0.259***	0.259***	0.260***	0.260***	0.260***	0.259***

	(0.00263)	(0.00263)	(0.00263)	(0.00264)	(0.00264)	(0.00264)	(0.00264)
Employment	0.0471***	0.0473***	0.0478***	0.0485***	0.0484***	0.0484***	0.0484***
	(0.00164)	(0.00164)	(0.00164)	(0.00165)	(0.00165)	(0.00165)	(0.00165)
Household size	0.00310***	0.00309***	0.00321***	0.00323***	0.00322***	0.00322***	0.00325***
	(0.000212)	(0.000212)	(0.000212)	(0.000213)	(0.000213)	(0.000213)	(0.000213)
White	0.295***	0.295***	0.295***	0.296***	0.295***	0.295***	0.293***
	(0.00234)	(0.00234)	(0.00234)	(0.00235)	(0.00235)	(0.00235)	(0.00236)
Urban		-0.00917***	-0.00933***	-0.00959***	-0.00910***	-0.00899***	-0.00887***
		(0.00173)	(0.00173)	(0.00174)	(0.00174)	(0.00175)	(0.00175)
Illness			0.0192***	0.0173***	0.0174***	0.0174***	0.0177***
			(0.00171)	(0.00175)	(0.00175)	(0.00175)	(0.00175)
Any chronic disease				0.00931***	0.00947***	0.00947***	0.00945***
				(0.00193)	(0.00193)	(0.00193)	(0.00193)
Health care services acceptable (PSU level), all responses					0.0162***	0.0163***	
					(0.00176)	(0.00177)	
Healthcare services acceptable (PSU level) – income quintiles 2, 3 and 4							0.0237***
							(0.00171)
Long distance to public healthcare facility (>=30 minutes, PSU level)						0.00123	0.000526
						(0.00243)	(0.00243)
Year controls (2009 as reference year)	YES	YES	YES	YES	YES	YES	YES
Constant	-0.600***	-0.600***	-0.605***	-0.611***	-0.626***	-0.627***	-0.632***
	(0.00505)	(0.00505)	(0.00507)	(0.00518)	(0.00544)	(0.00548)	(0.00544)
Observations	272,775	272,775	272,645	270,291	270,291	270,291	270,291
R-squared	0.371	0.371	0.371	0.372	0.373	0.373	0.373

Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

The health status variables, illness and chronic disease, both display positive relationships with medical scheme cover. Contrary to expectations, living in an urban area is negatively associated with the probability of medical scheme cover. The urban dummy does not distinguish between formal and informal urban areas and this may drive some of the negative association. Employment (in the broad sense) increases the likelihood of medical scheme cover by at least 4.7 percentage points. This relatively small association is most likely the result of the broad variable definition and the fact that expenditure and other variables (e.g. race) that are most likely associated with formal, skilled employment are already controlled for.

I expected a negative relationship between medical scheme membership and the variable controlling for perceived public healthcare acceptability, with medical scheme membership more likely in areas (PSUs) where public healthcare services are perceived to be less acceptable. However, in the regressions a small positive and significant association³⁰ was found. The positive association is most likely due to confounding factors and missing variables. Although geography was controlled for through inclusion of an urban dummy variable, it is possible that medical scheme members reside and/or seek care in areas where there is greater availability of both private and higher quality public healthcare services.

Reporting travel time of 30 minutes or more to the public healthcare facility typically consulted (at PSU level) has a very small positive and insignificant association with the likelihood of medical scheme cover. This may indicate that distance to public healthcare facilities is no longer a major healthcare access barrier influencing medical scheme choice. There has been a large-scale expansion of especially public primary healthcare facilities in South Africa since the early 1990s and this is therefore unlikely to be an important correlate of medical scheme membership.

For the longer time period (2002-2011), I ran two specifications (3-4), adding a variable to control for reported illness or accidents in the period before the survey in specification 4. For the shorter time period (2002-2008) I also ran two specifications (1-2) using the same approach for the longer period. The size and signs of the coefficients for this period are very similar to those for the 2009-2011 period where variables overlap.

The results from the GHS regressions indicate that income, education (or tertiary education), employment and race are the correlates that matter most for medical scheme membership. These variables are all related to socio-economic status and privilege. It is likely that race has such a large

³⁰ The coefficient size almost doubles when public healthcare acceptability observations is excluded for the richest and poorest quintiles due to reporting bias concerns from the acceptability variable in the regression (specification 7). The sign, however, remains the same.

association with medical scheme membership due to insufficiently nuanced control variables in the GHS for types or categories of employment. However, the LFS results indicate the importance of public sector and also skilled employment.

The next section considers the potential for expansion of the medical schemes market by exploring various affordability scenarios.

2.5 A middle-market for voluntary private health insurance

In the context of the private health insurance system that serves less than a fifth of the South African population and the importance of socio-economic correlates such as income or expenditure (affordability) and education in determining membership affirmed by the preceding analysis, there have been recent attempts to extend medical scheme cover to parts of the so-called “missing middle” or the middle market³¹. These attempts have mainly been focused on increasing the affordability of medical schemes or medical scheme-like products by lowering their costs, often with a more limited benefit package. Examples of these attempts include the following:

- **Low-Income Medical Schemes (LIMS):** While there have been initiatives such as the market-wide Low-Income Medical Schemes (LIMS) investigation concluded in 2006 which considered the viability of creating lower cost medical schemes (McLeod, 2012), very few such products are offered by current medical schemes. Meanwhile, members of the middle market are required to pay hospital fees for hospitalisation in public hospitals if these households earn more than R6,000 per month.
- **Government Employees Medical Scheme (GEMS):** GEMS was launched in 2006 to provide health insurance cover to government employees. At the same time, the government provided generous subsidies to especially low-income government workers which would effectively allow cover for a family of up to five members at no additional cost to workers (Govender et al., 2013). The launch of GEMS, to some degree, provided a natural experiment for the likely changes in health behaviour and health outcomes that can be expected from NHI.
- **Low-cost hospital cash plans:** In the last few years, low-cost hospital cash plan insurance has experienced high take-up in the face of increasingly expensive medical schemes. These policies provide a fixed amount of cover for every day spent in hospital, after a minimum

³¹ I did not explicitly deal with the issue of the missing (potentially lower cost) supply side which would be required to service the middle market. I therefore refer to the “middle market” when describing the demand-side and use the term “missing market” when referring to supply side that would have to service the “middle market”.

number of days (typically two) of hospitalisation. Hospital cash plans operate in a grey (and currently under revision) regulatory space, with estimates that there are 1-1.5 million such active policies, covering 2.4 million lives (Childs and Erasmus, 2012). The high take-up demonstrates an unmet need for at least some financial protection by mainly middle and lower income South Africans.

The CMS, the medical scheme industry regulator, recently released an official communication in the form of a circular requesting the medical schemes industry to provide input into its proposal on a “low cost benefit option” for the market (Council for Medical Schemes, 2015)³². According to the circular (Council for Medical Schemes, 2015: 2), “tentative market projections estimate that the cost of providing a low cost benefit option may vary between R200-R400 per month per beneficiary”. The document sets out a possible framework for the more limited benefits that such a product could provide and invites stakeholders in the medical schemes industry to provide input on the costing and benefits of the product. From the document, it seems that this type of product may have to provide more limited hospitalisation and medical specialist benefits.

Establishing an affordability threshold:

Evidence that affordability may be the binding constraint in increasing the size of the medical schemes market is provided by data on differential household coverage. According to the GHS of 2011, 82.2% of white households with at least one medical scheme member had medical scheme cover for all household members. This was the case for only 38.8% of black households where there was at least one medical scheme member present. Although differential coverage decisions could be driven by health risk, South Africa has a generally high disease and health risk burden and faces a so-called “quadruple” disease burden³³ (Mayosi et al., 2009). If this had been lower, it is likely that adverse selection would have played a much larger role. Households may also elect not to obtain coverage or coverage for certain members if the private healthcare services do not offer value in terms of greater healthcare acceptability compared to the public sector. Evidence from the GHS, however, indicate lower levels of healthcare acceptability in the public sector as signalled by reported long waiting times, rude staff and drug stock-outs (Burger et al., 2012).

³² After submitting this thesis for examination in October 2015, it was reported that the CMS withdrew the circular and proposed guidelines for the low-cost benefit option (Khan, 2015). According to media reports, this took place due to “to criticism from key stakeholders such as the South African Medical Association” (Khan, 2015).

³³ The “quadruple” disease burden includes communicable diseases such as HIV/AIDS and TB, non-communicable or lifestyle diseases, accidents (also due to high levels of violence) and maternal and child health problems.

Apart from affordability, differences in coverage by race also point towards differences in both household size and household structure (i.e. who lives in the household) between race groups. The nature of the household survey data does not allow for calculation of accurate medical scheme dependent ratios, as it is not possible to know which households members were covered by other members' insurance. In the context of increasing concerns about the affordability of medical scheme cover (Fish and Ramjee, 2007; Eighty20, 2009) and annual cost increases of about 4% in excess of the average inflation rate (Council for Medical Schemes, 2014), medical scheme dependent ratios as captured in administrative data have been decreasing. By 2013, the average dependent ratio was 1.3 (Council for Medical Schemes, 2014). According to McLeod and Ramjee (2007), affordability concerns may lead households to elect to provide cover for only certain members and, in particular, to deregister children from cover to minimise costs.

In exploring the affordability of medical schemes in South Africa, Fish and Ramjee (2007) test four levels of expenditure on medical schemes relative to monthly per capita income (5%, 10%, 15% and 20%,) but argue that a realistic level most likely lies between 5% and 10% of monthly per capita income. Using data from a large insurer's healthcare survey, they assume an employer contribution of 50% in calculating affordability for all four affordability thresholds (Fish and Ramjee, 2007). Eighty20 (2009) uses data from the Income and Expenditure Survey (IES) of 2006 to inform a choice of an affordability threshold of 10% - according to the IES of that year 70% of households that had any form of medical scheme expenditure in that year spent 10% or less of total household expenditure on membership.

I used a mid-point value (R300) in the R200-R400 price range suggested by the CMS (2015) for a possible low-cost benefit option to estimate the potential additional market that could be covered by this type of product. The CMS derive their price range from earlier research on affordability of medical schemes such as the Low-Income Medical Schemes (LIMS) initiative. In setting out its vision for more affordable medical schemes-like type products, the CMS (2015) provided two examples of a more limited benefit range for low-cost benefits options and invited proposals and submission from the medical schemes industry for costing these benefit options.

Two affordability scenarios informed by the work of Fish and Ramjee (2007) and Eighty20 (2009) are used in the market expansion calculations: 1) a 5% affordability threshold; 2) a 10% affordability threshold. Affordability in both scenarios is dependent on a 50% employer subsidy and the analysis is therefore focused on employed individuals and their likely dependents who reside in households that are partially covered or have no medical scheme coverage. I used data from the GHS of 2011, as this is the most recent data in the pooled dataset. Given that income data are often under-reported

in surveys, per capita household expenditure data which were adjusted for inflation purposes³⁴ were used. The price point was also adjusted for inflation purposes.

Using the above parameters, the estimates indicate the market could be extended to between 2.1 million (10% affordability scenario) and 5.3 million (5% affordability scenario) currently uncovered employed individuals, if an employer subsidy of 50% is assumed. This is a conservative estimate, as it is focused only on employed individuals and does not take into account the possibility of extending cover to family members or other beneficiaries. The upper bound estimate would imply growth of more than 60% in the current medical schemes market, increasing individuals who have some type of private health insurance cover to 27% of the South African population. If only a small percentage of these individuals choose to also extend their cover to family members, 30% or more of the South African population could be covered by private health insurance. It is, however, important to keep in mind that these estimates pertain only to healthcare financing and do not take into account the ability of the private healthcare market to effectively absorb additional demand for private healthcare services.

2.6 Conclusion

In this paper I explored how medical scheme membership changed over the period 2000 to 2011, what the regulatory barriers to market expansion are and what the correlates of medical scheme cover indicate about the potential for market growth. The descriptive analysis considered indicators of affordability and health status or risk. While I controlled for a number of socio-economic and demographic variables, the empirical analysis focused on three main categories of medical scheme demand correlates: employment and income (relating to affordability), health status (or risk) and user acceptability of healthcare. With the available data, I found that although health status or risk seems to matter, affordability as captured by socio-economic variables seems to matter more. Acceptable (or more acceptable) *public* healthcare services were positively associated with medical scheme membership which points towards the presence of confounding factors.

While there has been some growth in the medical schemes market in recent years, this was largely driven by the extension of medical scheme cover in the public sector. This process has now likely reached its end. Black South Africans as proportion of medical scheme members has experienced large growth but this cover is still limited to less than 10% of black South Africans.

³⁴ The Statistics South Africa Consumer Price Index (CPI) headline index numbers with December 2012 as baseline month (CPI in this month takes on a value of 100) were used.

The linear probability analysis using GHS data showed that race remains a large predictor of the likelihood of medical scheme cover after controlling for various socio-economic correlates and that white South Africans are significantly more likely to have medical scheme cover than their black counterparts. This most likely indicates that the other control variables did not sufficiently measure socio-economic status or points towards omitted variables. Since many black middle class members are likely to be first-generation middle class members they may be spending proportionally more of their income on assets in order to eliminate an asset deficit compared to other middle class members. This would mean that there would be proportionally less income available for expenditure on financial products such as medical scheme membership.

However, it was not possible to explore this further with the acceptability of public healthcare services variable, given that it most probably suffers from endogeneity due to the presence of confounding factors.

Three main barriers remain: the current nature of mandatory minimum health benefits, both in terms of definitions and implementation of the definitions, the lack of a risk equalisation method between medical schemes and voluntary rather than compulsory coverage for the employed which limits the size of the risk pool and could promote adverse selection. The latter two imply that the principal of social solidarity has not yet been fully implemented in the medical schemes market. If at least the first two barriers could be addressed and membership costs decreased due to these and other changes, the market could be extended to as much as 27% of the South African population. This could assist to improve the available per capita funding and functioning of the public healthcare system by helping to limit congestion and explicit rationing in the public healthcare system. Costs could also be limited through the implementation of a stricter referral system with general practitioners playing the role of gatekeepers to the services of medical specialists.

Limitations:

In considering the demand response of the extension of insurance to more South Africans it is important to recognise the limits of an analysis that simply measured associations between variables. In order to better predict or understand likely health seeking behaviour and demand in a UHC system as envisaged through NHI, causal analyses will ultimately be required. While the analysis considered the potential of market extension of private health insurance on the demand-side, the ability of the private sector to service this additional demand was not considered. This will be crucial to the success of any such market expansion and is quite a large limitation of this analysis

Implications for future research:

The limitations above also inform future research questions and approaches. It will be important to incorporate supply-side considerations and constraints more explicitly in an analysis such as this which considers the demand- and regulatory components of the medical schemes market. In particular, human resource constraints and the impact of alleged anti-competitive market behaviour a currently being investigated by the Competition Commission's Private Healthcare will have to be considered. Furthermore, a causal analysis or semi-causal analysis to consider socio-economic correlates of medical scheme cover is required. The construction of quasi-panels to track individuals' medical scheme selection characteristics and decisions over time could be one approach to create a better understanding of demand characteristics or determinants of medical scheme utilisation in South Africa.

Appendix A to Chapter 2

Table 2.A1: Number of medical scheme main members and members from the combined Labour Force Surveys (LFS) and Quarterly Labour Force Surveys (QLFS), the Council for Medical Schemes and the General Household Surveys (GHS)

Source: Various surveys and administrative data

Year	LFS* (2000-2007) and QLFS* (2008-2011) (Column A)	CMS medical scheme members* (Column B)	GHS (Column C)	CMS medical scheme beneficiaries (Column D)
2000	2,739,167	2,718,301	-	7,004,636
2001	2,744,551	2,740,572	-	7,025,262
2002	3,397,083	2,762,392	6,883,323	6,963,189
2003	3,289,203	2,802,815	6,914,804	6,924,686
2004	3,194,747	2,833,322	6,883,725	6,915,666
2005	2,709,580	2,812,083	6,548,722	6,835,621
2006	2,758,400	2,985,350	6,482,640	7,127,343
2007	3,119,865	3,233,490	6,817,517	7,605,236
2008	3,280,759	3,388,582	7,725,628	7,874,826
2009	3,335,213	3,488,009	8,286,191	8,068,505
2010	3,559,055	3,612,062	8,737,797	8,315,718
2011	3,526,707	3,730,565	8,050,934	8,526,409

Table 2.A2: Total medical scheme coverage by race

Source: General Household Survey, 2002-2011

Year	Total black cover	Total coloured cover	Total Indian cover	Total white cover
2002	2,884,583	754,166	325,221	2,919,353
2003	2,961,277	799,373	398,121	2,756,033
2004	2,662,157	752,879	410,351	3,058,338
2005	2,620,160	749,623	373,836	2,805,103
2006	2,718,192	675,205	339,045	2,750,198
2007	2,782,495	798,691	365,738	2,870,593
2008	3,220,111	941,086	487,026	3,077,405
2009	3,499,852	941,131	542,074	3,303,134
2010	4,040,810	970,821	608,086	3,115,080
2011	3,529,959	914,642	538,297	3,068,036

Table 2.A3: Black employment and medical scheme cover*Source:* Labour Force Survey and Quarterly Labour Force Survey, September 2000-2011

Year	Black formal sector employees	Black public sector employees	Black employees with cover	Black public sector employees with cover	Black private sector employees with cover
2000	4,093,213	1,171,555	1,082,301	557,772	524,529
2001	3,948,157	1,082,434	1,179,693	541,377	638,316
2002	4,079,575	1,059,075	1,698,953	750,077	948,876
2003	4,188,531	1,070,016	1,575,938	734,136	841,802
2004	4,485,638	1,060,415	1,430,469	684,507	745,962
2005	4,739,737	1,116,400	1,221,203	597,052	624,151
2006	5,075,698	1,147,161	1,264,598	607,350	657,248
2007	5,670,493	1,325,646	1,545,731	755,180	790,551
2008	5,913,554	1,287,963	1,799,808	897,857	901,951
2009	5,668,963	1,296,069	1,787,165	966,164	821,001
2010	5,636,035	1,310,419	1,922,281	1,003,217	919,064
2011	6,049,111	1,382,120	1,974,918	1,013,164	961,754

Table 2.A4: White employment and medical scheme cover*Source:* Labour Force Survey and Quarterly Labour Force Survey, September 2000-2011

Year	White formal sector employees	White public sector employees	White employees with cover	White public sector employees with cover	White private sector employees with cover
2000	1,792,377	331,969	1,192,503	272,491	919,018
2001	1,845,376	282,585	1,053,005	211,090	841,915
2002	1,813,731	260,050	1,119,272	198,049	921,223
2003	1,871,835	286,900	1,152,706	225,341	927,365
2004	1,836,961	302,378	1,153,368	253,370	898,847
2005	1,803,905	291,235	1,001,457	217,130	784,327
2006	1,819,208	235,871	938,875	173,865	765,010
2007	1,889,466	215,233	1,081,955	184,945	897,010
2008	2,070,651	218,209	913,847	190,169	723,678
2009	1,995,814	239,709	927,432	208,471	718,961
2010	2,008,718	234,569	957,800	211,580	746,220
2011	2,002,586	218,493	915,359	201,124	714,235

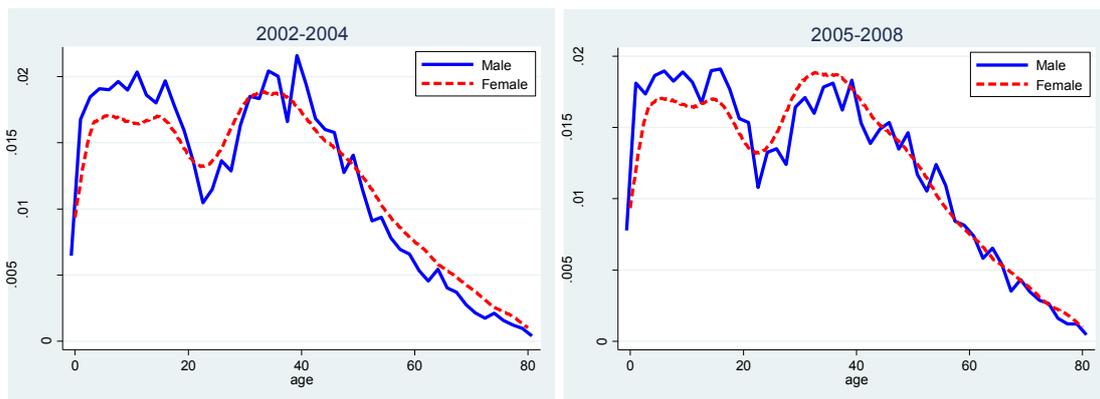
Table 2.A5: Reported chronic disease and illness (or injury) amongst those with medical scheme cover and those without

Source: General Household Surveys, 2009-2011

	Medical scheme cover n (%)	Uninsured n (%)	P-value
<i>One or more chronic diseases reported:</i>			
2009	1,407,870 (17.11)	4,940,696 (12.26)	0.000
2010	1,417,696 (12.17)	4,921,923 (16.37)	0.000
2011	1,473,805 (18.41)	5,164,586 (12.40)	0.000
<i>Illness or injury in 6 weeks preceding survey:</i>			
2009	1,909,512 (23.04)	7,142,427 (17.55)	0.000
2010	1,198,306 (13.73)	4,357,817 (10.68)	0.000
2011	1,077,588 (8.88)	3,723,184 (13.40)	0.000

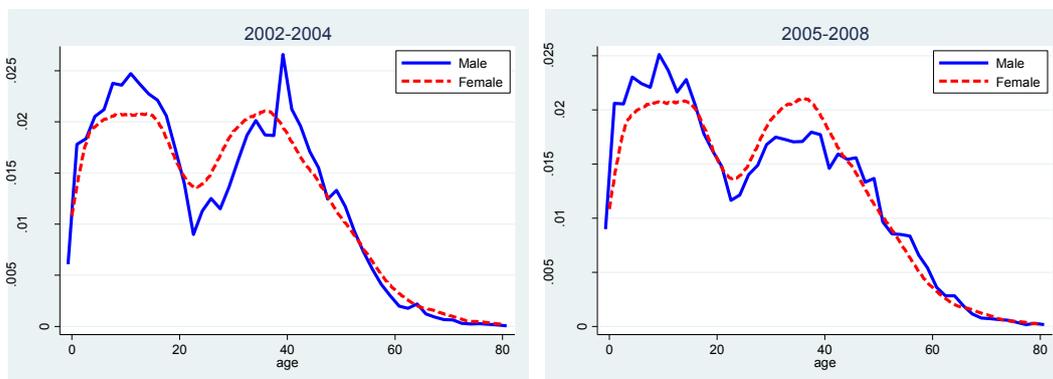
Figures 2.A1: Kernel density function of age distribution of all medical scheme members for two time periods: 2002-2004, 2005-2008

Source: General Household Surveys, 2002-2008



Figures 2.A2: Kernel density function of age distribution of all black medical scheme members for two time periods: 2002-2004, 2005-2008

Source: General Household Surveys, 2002-2008



Figures 2.A3: Kernel density function of age distribution of white medical scheme members for two time periods: 2002-2004, 2005-2008

Source: General Household Surveys, 2002-2008

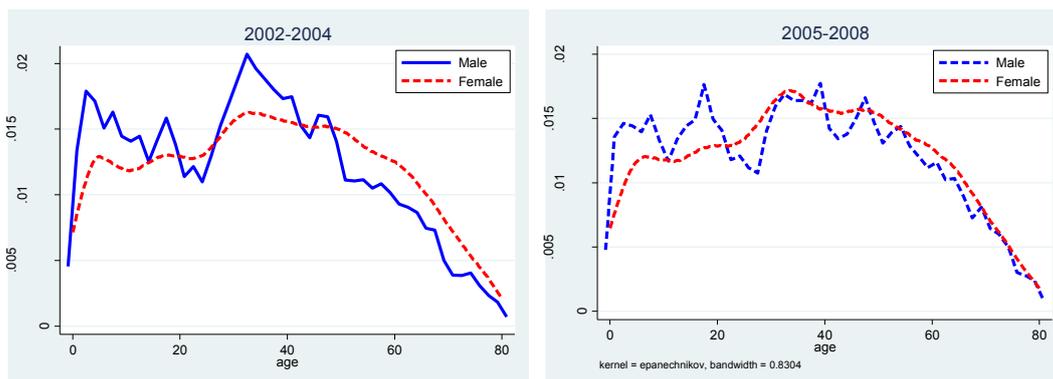


Table 2.A6: Output of linear probability model using GHS: 2002-2005, 2002-2011

Variables	Dependent variable: medical scheme cover			
	2000-2008		2002-2011	
	(1)	(2)	(3)	(4)
Log of per capita household expenditure	0.0935*** (0.000414)	0.0932*** (0.000414)	0.0958*** (0.000349)	0.0955*** (0.000349)
Formal housing	0.0356*** (0.000868)	0.0356*** (0.000868)	0.0367*** (0.000758)	0.0367*** (0.000758)
Clean water access	-0.0121*** (0.00110)	-0.0123*** (0.00110)	-0.0109*** (0.000983)	-0.0112*** (0.000983)
Flush toilet	0.0351*** (0.000861)	0.0352*** (0.000861)	0.0354*** (0.000736)	0.0355*** (0.000736)
Male	-0.0128*** (0.000697)	-0.0123*** (0.000697)	-0.0128*** (0.000598)	-0.0123*** (0.000598)
Married	0.0544*** (0.000945)	0.0546*** (0.000945)	0.0561*** (0.000815)	0.0563*** (0.000815)
Age 0-4	0.0836*** (0.00177)	0.0867*** (0.00178)	0.0812*** (0.00152)	0.0836*** (0.00152)
Age 5-14	0.0899*** (0.00158)	0.0949*** (0.00159)	0.0866*** (0.00136)	0.0905*** (0.00137)
Age 15-20	0.0707*** (0.00165)	0.0757*** (0.00166)	0.0677*** (0.00143)	0.0719*** (0.00143)
Age 21-40	0.0189*** (0.00146)	0.0227*** (0.00146)	0.0130*** (0.00126)	0.0162*** (0.00127)
Age 41-60	0.0318*** (0.00158)	0.0332*** (0.00158)	0.0276*** (0.00137)	0.0288*** (0.00137)
Household education 15 yrs plus	0.232*** (0.00160)	0.232*** (0.00160)	0.241*** (0.00137)	0.241*** (0.00137)
Employed	0.0388*** (0.000980)	0.0396*** (0.000981)	0.0416*** (0.000844)	0.0423*** (0.000844)
Household size	0.00587*** (0.000134)	0.00604*** (0.000135)	0.00495*** (0.000114)	0.00511*** (0.000114)
White	0.287*** (0.00142)	0.288*** (0.00142)	0.289*** (0.00122)	0.290*** (0.00122)
Illness		0.0279*** (0.00107)		0.0250*** (0.000909)
Year controls (2002 as reference year)	YES	YES	YES	YES
Constant	-0.548*** (0.00304)	-0.555*** (0.00305)	-0.556*** (0.00265)	-0.562*** (0.00265)
Observations	664,652	664,318	937,427	936,963
R-squared	0.364	0.365	0.366	0.367

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Appendix B to Chapter 2

Regulation of the medical schemes market

Van den Heever (2012) identifies three phases of regulatory development in the medical schemes market in South Africa:

- **The start of the medical schemes market with the establishment of occupational and industry schemes (approximately 1889-1988):** The first occupational health fund was established in the mining industry in 1889, leading to the establishment of several other such funds in the same industry. The phase was characterised by medical schemes being “non-competing, occupation funds sponsored by employers or industries” and mainly limited to the white population (Van den Heever, 2012: S5). The first real regulatory framework was introduced in 1967 and attempted to locate legal definitions spread throughout different pieces of legislation up to that point in one framework to enable registration and compliance. Van den Heever (2012) describes contributions as being differentiated by income during this period but notes that differentiation on health status was prohibited.
- **Rise of commercialism in the medical scheme sector and deregulation of the market during the 1990s (approximately 1989-1999):** Both the development of a private hospital market and the expansion of medical scheme administrators into the provision of health insurance during the 1980s led to commercialism as operational principle being embedded in the medical schemes market. In 1989 the government implemented legislation which allowed for premium setting based on the risk profiles of members and a further move away from social solidarity and social protection occurred when in 1994 the requirement for mandatory minimum benefits were dispensed with. These changes led to large industry movements away from occupational schemes to open schemes, the latter competing on the basis of risk-selection (selection of beneficiaries with the best risk profiles) and risk-rating (tailoring premiums relatively to risk profiles of members). According to Van den Heever (2012: S5) during this phase “regulation involved little more than the registration of medical schemes and some low-key prudential supervision”.
- **Medical scheme re-regulation (2000 plus):** The last and also current phase of regulatory development was an attempt to reintroduce the principles of social solidarity present in the first phase of development. In 1998 a new medical schemes act, the Medical Schemes Act No. 131 of 1998, was implemented, only taking effect in 2000. The new act established a more comprehensive regulatory framework and reintroduced three main principles to allow

for social solidarity: open enrolment, community rating and mandatory minimum benefits. However, it stopped short of mandating medical scheme membership for the population or certain sub-groups and also did not require risk equalisation or differentiation of premiums based on income.

Imperfect implementation of social solidarity:

Risk subsidisation was originally envisaged to occur through the implementation of a central risk equalisation fund (REF), to follow after the new medical schemes act which would transfer funds between medical schemes based on the risk profiles of scheme members, i.e. funds with members with better risk profiles would help subsidise funds with members with worse risk profiles. It is argued that this would have allowed schemes to compete on cost-effective delivery rather than risk-selection (McLeod and Grobler, 2010). While medical schemes did participate in a test phase of preparing and submitting mock REF submissions to the CMS after 2005 and a risk equalisation formula was developed, the REF was not legislated (McLeod and Grobler, 2010) and therefore not implemented (Ramjee et al., 2014). At least two other aspects which would have ensured achievement of social solidarity were also never implemented although proposed: income cross-subsidies and compulsory membership for at least the formally employed (Van den Heever, 2012). McLeod and Grobler (2010) argue that if the REF had been implemented before moving to a compulsory medical scheme system and the implementation of income cross-subsidies, members may have experienced large cost increases which would have led to a loss of members. There is thus a very specific sequence in which social solidarity reforms have to be implemented (McLeod and Grobler, 2010).

The cost implications of Prescribed Minimum Benefits (PMBs):

Prescribed Minimum Benefits (PMBs) were implemented to protect members by providing an assured minimum level of financial protection. PMBs include a list of 25 chronic diseases, 270 medical conditions and any emergency condition which medical schemes are required to cover at full cost (Republic of South Africa, 1999). Ramjee et al. (2014: 99) assert that “there is... strong contention from stakeholders that PMBs increase the contribution rates for medical schemes, due to both poor benefit definitions and unregulated medical prices”. Much of the critique is aimed at the fact that the definitions of PMBs, or arguably the implementation of the definitions, favours hospitalisation (curative) over preventative care and also leads to a bypassing of lower levels of medical workers such as general practitioners (GPs) in favour of medical specialists (Ramjee et al, 2014). Analysis indicates that PMBs are responsible for as

much as 60% of medical schemes actuarial liability (Ramjee et al. 2014). In the absence of a more flexible or less expensive approach to implementation of the PMB definitions, PMBs essentially determine the lower limit of medical scheme costs. It has been asserted that part of the high cost of the implementation of PMBs has been unregulated or uncapped medical prices (Taylor et al., 2007), the issue of which is currently being investigated through an enquiry of the Competition Commission in the private healthcare and funding market.

Chapter 3

Does insurance affect healthcare utilisation and provider choice in a polarised healthcare market? Evidence from a South African natural experiment

3.1 Introduction

The demand for health services is notoriously difficult to disentangle from the demand for health insurance³⁵. In a voluntary health insurance environment, individuals may tend to self-select into health insurance based on self-assessed risk and expected utilisation, leading to an over-representation of individuals that are overly cautious or sick amongst the insured, which inflates health visits. This may induce an artificial positive correlation between health visits and insurance, which confounds the estimation of the causal effect of insurance using observational data.

In an effort to better understand what health benefits universal coverage may offer, a number of recent experimental studies have attempted to disentangle this relationship by allocating insurance or eligibility for insurance exogenously, often via a lottery, thus enabling causal estimations of how insurance affects health service utilisation and provider decisions (Card, Dobkin and Maestas, 2008; Finkelstein et al., 2012; Levine, Polimeni and Ramage, 2014; Thornton et al., 2010). These studies examine healthcare choices when insurance facilitates access to better health services and/or lowers the cost of existing services. In developing countries there is reliable and robust evidence that insurance influences provider choice and out-of-pocket expenditure, but with elusive public health benefits because there is frequently no impact on healthcare utilisation (e.g. King et al., 2010; Wagstaff, 2010; Thornton et al., 2010). The United States insurance market, however, is the exception. Experiments with insurance show a significant and sizeable impact on utilisation (Card, Dobkin and Maestas, 2008; Finkelstein et al., 2012). This may be attributable to the polarised American health system and the consequent large perceived gap between the *ex ante* scenario without insurance and *ex post* scenario with insurance.

Against the backdrop of this emergent international literature on the impact of health insurance, I consider the launch of the Government Employees Medical Scheme (GEMS) in South African in 2006 as a natural experiment in expanding insurance coverage in a polarised health system. Under GEMS all government workers became eligible for health insurance subsidies and low-earning employees

³⁵ This difficulty is typically noted by papers that attempt to untangle part of this complex relationship. See, for example, Dong (2013) and Meer and Rosen (2014).

received a full subsidy for the most basic benefit package, Sapphire. Under this scheme, no co-payments were required when using network providers. GEMS has had a dramatic impact on South Africa's medical schemes landscape. Between 2006 and 2012, while there was little growth in the rest of the medical schemes market, GEMS provided healthcare cover to 370,000 previously uninsured households (Moloabi, 2013). Since its launch GEMS has generated a steady stream of enrolments and by 2013 GEMS was still receiving 7,000 monthly applications, making it the fastest growing medical scheme in South Africa and the country's second largest (Moloabi, 2013).

The introduction of GEMS allows for the estimation of how the extension of healthcare coverage to uninsured public sector employees has impacted healthcare utilisation and provider choice, adding to the available international literature on this topic. The South African health system is an interesting case study to consider because of the polarised provider landscape which is not present in most of the countries in which this research question has been examined. Following the extreme contours of one of the world's most unequal societies, there are dramatic disparities between the health services available to the poor and the affluent. Private providers tend to charge prices that are prohibitively expensive for most South Africans and consequently only a small subsection of affluent individuals (often with comprehensive coverage) has reliable and frequent access to these providers. While the private sector represents 51.5% of South Africa's health expenditure (WHO, 2015a), only 17% of South Africans are medical scheme members (Council for Medical Schemes, 2014). By contrast, public sector providers are visited almost exclusively by the uninsured and less affluent segment of South Africa's population.

Although there is no representative evidence on the gap in clinical quality between the private and public health sector, there are glaring differences in access to nurses, doctors and specialist (McIntyre et al, 2007). GPs are usually the entry point into the private provider system, while nurses represent the entry point for the public system. Survey analysis shows that public sector facility visitors are considerably more likely to complain about rude staff, drug stock outs and long waiting times. (Burger et al., 2012). At private clinics waiting times ranged between 10 and 40 minutes versus 50 minutes to 3 hours at public sector clinics (Palmer, 2002). Consequently, it is not surprising that recent studies report substantial differences in the perceived quality of care offered by private and public providers. A discrete choice study on the nature of preferences for public healthcare in the Western Cape and Eastern Cape provinces of South Africa found "a general preference not to utilize public health facilities" and concluded that if government wanted to boost clinic visits, they would have to improve public health facilities (Honda et al., 2014: 9). Similarly, McIntyre et al

(2009:725) indicate that individuals were only willing to contribute to public health services if they were assured of the quality of such health services.

The purpose of this chapter is to add to the small but growing body of international literature on the impact of health insurance extension on health seeking behaviour. It is one of few studies that utilises a natural experiment and cross-sectional data to examine this question. It is also of pragmatic value to local planning needs.

In a South African context this question is relevant and pertinent because it provides some indication of the magnitude of the demand shifts that can be anticipated under the proposed NHI plan. The South African government released a policy proposal on the implementation of a NHI scheme in 2011 (Department of Health, 2011a).³⁶ In the NHI policy proposal, it is suggested that the NHI will contract in the services of general practitioners or doctors from the private sector (Department of Health, 2011a). Interpreting public sector employment as an exogenously administered expansion of eligibility for insurance, the estimates provide a reliable benchmark of how the proposed NHI may affect choices amongst providers, especially at the high volume entry level access points, i.e. public sector clinics and private sector GPs. Providing an estimate of the size of such shifts can help to guide the government's workforce planning and inform their rationing and gatekeeping strategies.

3.2 The relationship between insurance and healthcare utilisation: international evidence³⁷

The ultimate objective of health insurance is to provide access to appropriate healthcare services in order to improve health outcomes. There is presently no clear evidence on the health impact of health insurance, although increasingly studies that use experimental data or causal methods do find evidence of increased healthcare access or utilisation. While experimental methods allow for an analysis of the relationship between health insurance and healthcare access (also health outcomes) in a causal manner by overcoming complexities related to endogeneity and confounding factors, the methods themselves may lead to a sub-optimal maximisation of healthcare access. Randomised control trials on the impact of health insurance coverage often experience problems with low, or lower than expected, take-up of insurance offered at low(er) prices or for free in a voluntary setting

³⁶ If fully implemented, the NHI will provide universal health coverage to all South Africans. Although no further details on the financing component of the proposal have since been publically released, government officials have continued to refer to NHI as the government's position on the future financing of healthcare in South Africa (Gordan, 2014; Motsoaledi, 2014).

³⁷ Although a variety of electronic search engines (Google Scholar, Econlit, ScienceDirect) were used to identify the relevant literature covered in this section, as there are really only a handful of papers that use causal analysis (natural experiments and/or randomized control trials) to consider the relationship between insurance and healthcare utilisation, most of the covered literature were identified from the bibliographies of the most recent papers on the topic.

(e.g. Thornton et al., 2010; Finkelstein et al., 2012). Although this may relate to too short study periods, it could also be due to the quality signal provided by free or lower price healthcare providers. Although existing studies suffer from some of the drawbacks described above, it still provides valuable background. A short summary of some of the causal or experimental literature on the impact of health insurance on healthcare utilisation is therefore provided.

A randomised control trial on the impact of Seguro Popular in Mexico on utilisation and expenditure found no impact on utilisation or on diagnoses over a 10-month period but did find a reduction in out-of-pocket expenditure on health (King et al., 2010). The absence of an impact on utilisation was ascribed to the short study period. Similarly, health insurance cover for the poor in Vietnam successfully reduced out-of-pocket expenditure on health services but did not impact total utilisation (Wagstaff, 2010). In rural Cambodia, a randomisation study via a lottery which provided “winners” the option to buy micro-health insurance did not find an overall increase in utilisation, although insurance cover did lead to changes in provider choice (Levine et al, 2014). The authors identified an increase in the use of covered public health facilities for emergency health conditions, while the insured decreased their use of non-covered private healthcare and pharmacies (Levine et al, 2014). When social health insurance was offered to informal sector workers through an insurance lottery in Nicaragua (similar to the experiment in Cambodia), no impact on overall utilisation was identified, although the insured switched towards health services at covered facilities (a mix of public and private providers) and out-of-pocket expenditure decreased (Thornton et al., 2010).

Due to its polarised healthcare provider market, the United States healthcare market may be the most comparable to that of South Africa. Two recent experiments showed significant impact on utilisation rates. Finkelstein et al. (2012) offered insurance coverage to randomly selected members of a group of low-income uninsured adults in Oregon and found a notable and significant increase in out-patient and in-patient healthcare utilisation and a decrease in out-of-pocket healthcare expenditure and medical debt. Card, Dobkin and Maestas (2008) found that eligibility for Medicare at the age of 65 leads to an increase in healthcare utilisation. As expected, the increased use of lower cost services, e.g. doctor’s visits, was concentrated amongst the elderly who had the lowest rates of insurance coverage before Medicare eligibility, while increased use of higher cost and more elective type of procedures, e.g. bypass surgery and knee replacement, was concentrated amongst the elderly who were more likely to have held supplementary health insurance after Medicare eligibility.

Research in the Philippines show that where health insurance is offered on a voluntary basis, there is often a lag in the utilisation of the product due to a learning process, or understanding of the working of insurance, that has to take place. Investigating household health decisions when a young child is hospitalised, Quimbo et al. (2008) find significant under-utilisation of benefits by newly enrolled beneficiaries of PhilHealth. This effect was more pronounced at lower levels of maternal education. The authors argue that the underutilisation could be due to a lack of awareness of the benefits provided by PhilHealth. As will be discussed below, this could help explain why the initial take-up of GEMS was slower than expected, especially amongst lower salary employees who may have been insured before.

3.3 The launch of the Government Employee Medical Scheme (GEMS)

The 1999 Remuneration Policy Review by the South African government identified a number of concerns with health insurance provision, including “inequality in access to medical scheme cover, affordability concerns, lack of value for money, spending inefficiencies and little integration with public sector healthcare” (McLeod and Ramjee, 2007: 55). At the time, less than half of employees had health insurance cover (McLeod and Ramjee, 2007).³⁸

In response to these problems, a legal framework was drafted for the establishment of a government employees medical scheme in 2002 (McLeod and Ramjee, 2007), following which GEMS was legally registered in January 2005 and started to actively recruit members in January 2006 (GEMS, 2012).

To encourage government employees enrolled in open medical schemes to join the newly established health insurance scheme in the absence of a track record on good administration and coverage, benefit options were competitively priced³⁹ and the government offered increased subsidies to 75% (cf. 66% for open schemes) for workers enrolled in GEMS.⁴⁰

GEMS offered five benefit options: Sapphire, Beryl, Ruby, Emerald and Onyx. The Sapphire and Beryl schemes were targeted at lower income employees (McLeod and Ramjee, 2007). Sapphire beneficiaries were restricted to public hospitals and Beryl beneficiaries were offered access to a

³⁸ It is, however, important to note that despite the criticism against the previous public sector approach to health insurance, at the time GEMS was launched government employees were already receiving quite generous health insurance subsidies. Government provided a subsidy of up to two thirds of the cost of any health insurance scheme selected by staff members (McLeod and Ramjee, 2007).

³⁹ According to GEMS's (2011b) own comparisons, benefits less expensive than that provided by other schemes as it was provided at a discount of 10% to 25% compared other medical schemes.

⁴⁰ To limit the government's exposure to increasing insurance premiums, a cap was proposed for the subsidy (McLeod and Ramjee, 2007). Initially the cap was R2020 but by 2013 the subsidy limit for Sapphire had increased to R2760 for these workers (GEMS, 2013).

network of private hospitals. Ruby, Emerald and Onyx were more generous and covered hospitalisation in any private hospital (McLeod and Ramjee, 2007). The Ruby option included a personal medical savings account which could be used for day-to-day medical expenses, while the Emerald and Onyx options offered comprehensive health coverage for both private out-patient and in-patient providers and were mainly targeted at higher income earners. Emerald is the largest scheme and was designed to be comparable to health insurance packages offered by open schemes.

The government provided a 100% subsidy of the lowest tier benefit option to employees on salary level 1-5⁴¹ if they chose to become members of Sapphire, the lowest tier benefit option. In fact, the subsidy for these workers was sufficiently generous to allow a low-salary government worker and her family to obtain cover on the Sapphire option without any out-of-pocket contributions (McLeod and Ramjee, 2007).

Figure 3.1: Total lives (members and beneficiaries) covered by GEMS, 2006 – 2013

Source: Total members and beneficiaries on 31 December of each year from GEMS annual reports 2006-2013.

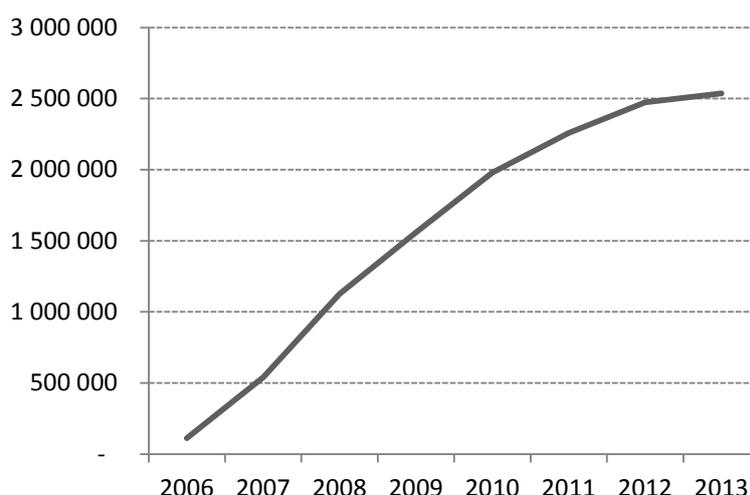


Figure 3.1 shows clear, steady growth in enrolment in GEMS between 2006 and 2013. By 31 December 2013, GEMS covered more than 2.5 million lives via 684,281 government employees (Council for Medical Schemes, 2014). These 2.5 million lives represented 5% of the South African population and more than a quarter of the South African medical schemes market. More than half of these members were previously uninsured (Moloabi, 2013). By the end of 2013, 60.3% of public sector employees were members of GEMS (GEMS, 2013).⁴² Health insurance coverage in the public

⁴¹ Equivalent to a threshold of R9,000 per month in 2013.

⁴² Health insurance coverage in the public sector is, however, likely to be higher than the 60.3% as employees who were previously members of other schemes were allowed to retain their membership, albeit foregoing the more generous subsidies for GEMS.

sector is, however, likely to be higher than 60.3% as employees who were previously members of other schemes were allowed to retain their membership, albeit foregoing the more generous subsidies for GEMS.

Unlike many other medical schemes, the majority of GEMS members are female, while the majority of beneficiaries (family members) are also women. The scheme argues that “this is to be expected, given the large percentage of female employees employed in the health and education sectors” (GEMS, 2012: 42). In 2011, almost 66% of GEMS members were women (GEMS, 2011).

Enrolments increased gradually, but with little evidence of bias based on geographical reach or salary bands. There were no precedence or priority treatment to any group in enrolment. The only exception is newly appointed staff who were not given the option of remaining or becoming members of external schemes but were required to join GEMS. Given the high levels of unemployment in South Africa amongst the unskilled and the prevalence of comprehensive insurance amongst skilled formal sector employees, membership of new appointments is unlikely to be a significant concern.

A survey of educational and health sector government employees confirmed previous findings on constraints to insurance enrolment by identifying affordability, the perceived administrative complexity of registration and lack of access to information about benefit options as constraints (Govender et al., 2013). The study shows that individuals with lower levels of education were less likely to enrol (Govender et al., 2013).

3.4 Methodology

The econometric model in this chapter considers health seeking behaviour, y , of individual i at period t . This is expressed as a function of whether or not the individual has health insurance (in which case $h = 1$, otherwise $h = 0$), as well as other observable (\mathbf{x}) and unobservable (u) determinants:

$$y_{it} = \alpha h_{it} + \mathbf{x}_{it}\boldsymbol{\beta} + u_{it} \quad [1]$$

The interest of the econometric analysis is primarily in identifying and estimating the causal effect of insurance on health seeking behaviour, represented by the treatment effect parameter α .

Medical scheme members effectively face an altered set of health service prices due to medical scheme reimbursements but also subsidies and tax benefits associated with insurance. However, identifying this effect using observational data is complicated by the fact that health insurance is not randomly administered to individuals, but is rather the outcome of an expected cost-benefit analysis

based on a variety of considerations many of which are unobserved by the econometrician. It can be expected that individuals with more resources, who place greater importance on good health or with a history of health problems, may be more inclined to seek expensive healthcare regardless of whether insured or not. However, these are the same individuals who would have more to gain from medical insurance coverage. A positive correlation between insurance and medical visits is therefore not necessarily indicative of the causal effect of insurance on healthcare choice. Controlling for measures of perceived health, household income and other covariates should partly address this problem, but some upward bias in the estimate of α is likely to remain due to the presence of unobservable resource constraints and preference factors. On the other hand, if health insurance is measured with error then its effect on behaviour may suffer from attenuation bias.

An instrumental variable (IV) strategy can identify the causal impact of insurance, but convincing IVs are notoriously rare. According to Angrist and Pischke (2008: 86) “good instruments” derive from “institutional knowledge and ideas about the process which generates outcomes”. In the case of this analysis, an IV is required to affect the inclination to have medical insurance without directly affecting health seeking behaviour. Such an IV is provided by the introduction of GEMS. In 2006 all government employees became eligible for insurance subsidies, the subsidy was increased and there were significant efforts to provide attractive health insurance coverage options for employees with low salaries. Crucially, GEMS only affected the behaviour of public sector employees and their households, which suggests that other workers may – after making the necessary adjustment for differences in composition – provide a useful counterfactual for behaviour and choices in the absence of this scheme.

The proposed identification strategy combines elements of the difference-in-difference and IV estimators (see Angrist and Pischke, 2008). In the first-stage regression a difference-in-difference approach is used that allows for extraction of the exogenous variation in health insurance due to the implementation of GEMS. Suppose health insurance is determined according to the following process

$$h_{it} = \mathbf{x}_{it}\boldsymbol{\theta} + \pi p_{it} + \delta_t + \gamma_t p_{it} + e_{it} \quad [2]$$

where p_{it} is a public sector dummy variable ($p_{it} = 1$ if the individual is a government employee, and 0 otherwise), δ_t represents the period t health insurance effect and γ_t is the additional post-GEMS government employee effect. The public sector effect, π , controls for the fact that public sector employees may be different from other workers in unobservable ways that affect the health insurance choice, while the time effects δ_t represent economy-wide changes in medical costs, health awareness and other unobservable determinants of the decision to obtain health insurance. Finally,

the γ_t coefficients are defined so that $\gamma_t = 0$ if $t < \tau$, where τ is the date of the implementation of GEMS. Equation [2] is estimated by regressing medical scheme coverage on the observable covariates, a set of time dummies, a public sector dummy and a public sector dummy interacted with time dummies for all periods since the implementation of GEMS in 2006. It is worth noting that the use of time dummies allows for a more general time trend than a specification that includes a linear time trend only. Furthermore, one can explicitly test the validity of the assumption that public sector workers have a similar pre-GEMS time trend as the rest of the economy by interacting the public sector dummy with all the period dummies and formally testing whether $\gamma_1 = \dots = \gamma_{\tau-1} = 0$.

The coefficients on the time-government sector interactions represent the difference-in-difference first-stage effects of interest. Under the assumption that $E(e_{it}|p_{it}, t, \mathbf{x}_{it}) = 0$ it follows that

$$\gamma_\tau = \{E(h_{it}|p_{it} = 1, t = \tau, \mathbf{x}_{it}) - E(h_{it}|p_{it} = 0, t = \tau, \mathbf{x}_{it})\} \\ - \{E(h_{it}|p_{it} = 1, t = \tau - 1, \mathbf{x}_{it}) - E(h_{it}|p_{it} = 0, t = \tau - 1, \mathbf{x}_{it})\}$$

and similarly for γ_t where $t > \tau$. Intuitively, if the time trend in health insurance coverage for those not in the public sector provides an accurate counterfactual for how government employees would have behaved in the absence of GEMS, then γ_t represents the causal effect of GEMS on the probability of having health insurance in period $t > \tau$ for government employees. These effects can be consistently estimated using OLS, and the estimates used to construct a predicted health insurance membership variable \hat{h}_{it} which is purged of all endogenous variation. Rewriting equation [1] to include public sector and time dummies, and replacing health insurance with its predicted value produces the following reduced form equation:

$$y_{it} = \alpha \hat{h}_{it} + \mathbf{x}_{it}\boldsymbol{\beta} + \mu p_{it} + \omega_t + u_{it} \quad [3]$$

A least squares regression of equation [3] will now provide a consistent estimate of the treatment effect, as long as $E(u_{it}|p_{it}, t, \mathbf{x}_{it}, \hat{h}_{it}) = 0$. The crucial restriction is that variation obtained from the IVs must be uncorrelated to the unobservable determinants of healthcare choice after conditioning on period, public sector employment and the vector of observable covariates \mathbf{x}_{it} . This is analogous to the assumption of a common trend in health insurance coverage for public sector employees and the rest of the population in the absence of GEMS.

3.5 Data

The estimation strategy outlined above requires information on health seeking behaviour, health insurance and the industry of employment from 2002 to 2012. Unfortunately, no dataset meets all the aforementioned information requirements so two alternative estimation strategies are pursued. The first utilises the reliable health insurance and industry information in the LFSs (the biannual LFS,

from 2000 to 2007, and the QLFS, since 2008)⁴³ and the useful and detailed information on health services in the GHS⁴⁴. This approach is required because the LFSs do not have any information on health seeking behaviour and the GHSs under-capture medical scheme membership during the crucial first years of GEMS (Figure 3.2). These two datasets are then combined to produce a two-sample 2SLS estimate of the causal effect of interest. The first-stage estimates equation [2] using the LFS/QLFS data, after which these estimates are applied to the GHS data in order to estimate equation [3] with the instrumented insurance variable.

The second strategy uses the three waves (2008, 2010, 2012) of the National Income Dynamics Study (NIDS) to estimate the first and the second stage of the instrumented variable regression. NIDS is a nationally representative panel survey covering about 7,000 households. Because the first wave occurs in 2008, the data does not allow comparison with the pre-GEMS period and I therefore use the gradual GEMS roll out (see Figure 3.1) and consequent increase in the likelihood of being a member of a medical scheme amongst public sector workers over the four-year window to capture the exogenous impact of GEMS.

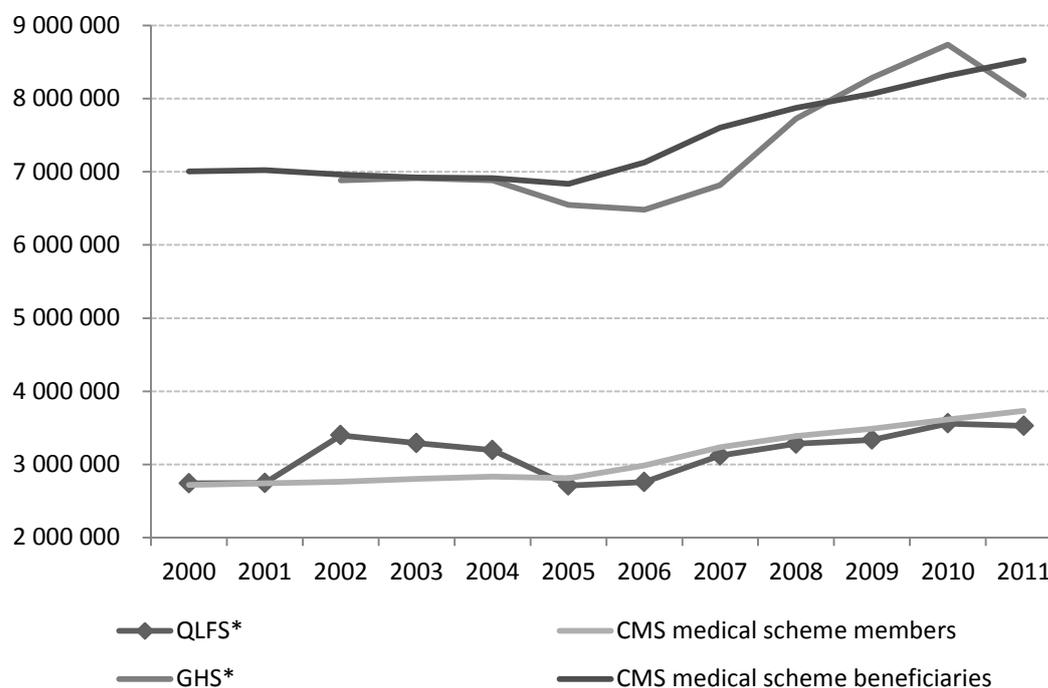
Figure 3.2, below, considers levels and time trends for beneficiaries and members, comparing the administrative data from the Council for Medical Schemes with estimates from the GHS and the LFS. It is encouraging to see that the survey data tracks the administrative data remarkably well – the only exceptions being a slight LFS over-count between 2002 and 2004 and a GHS undercount between 2006 and 2008.

⁴³ The LFS/QLFS collects information on the labour market status and activities of a sample of individuals living in South Africa who are older than 14. It covers about 30,000 dwellings.

⁴⁴ The GHS is an annual survey and aims to gather information on the circumstances and quality of life of households. It includes approximately 25,000 households.

Figure 3.2: Growth in medical scheme members and beneficiaries, 2002-2011

Source: GHS and QLFS 2002-2011, CMS 2002-2011



Although the empirical strategy outlined in Section 3.4 requires identifying public sector workers from the survey data, both the GHS and the NIDS data only provide industry of employment information at the one-digit ISO level. The LFS data (which provide more detailed industry data, as well as sector of employment) reveal that there is substantial overlap between public sector employment and working in the Community and Social Services (CSS) industry⁴⁵. This suggests using being employed in this industry as a proxy for government employment. Of course, this adds measurement error to the IVs, which may affect the estimates.

Although there are some (11%) public sector workers who work outside of this industry, the more serious issue is that only 63% of workers in this industry are public sector workers. The identification strategy assumes that there is a common time trend in health insurance and health seeking behaviour that affects all non-government workers, including the non-government workers in the CSS industry. This implies that the OLS estimates of coefficient γ_2 in equation [2] will be attenuated by 0.63, since this coefficient is the weighted average of the true GEMS effect on government employees and zero on the remainder of the industry. Similarly, regressing health seeking behaviour on the interaction between the CSS industry and the post-GEMS periods will suffer from the same attenuation rate. Since the 2SLS estimator can be calculated as the ratio of these two effects, the

⁴⁵ This includes the health and teaching professions.

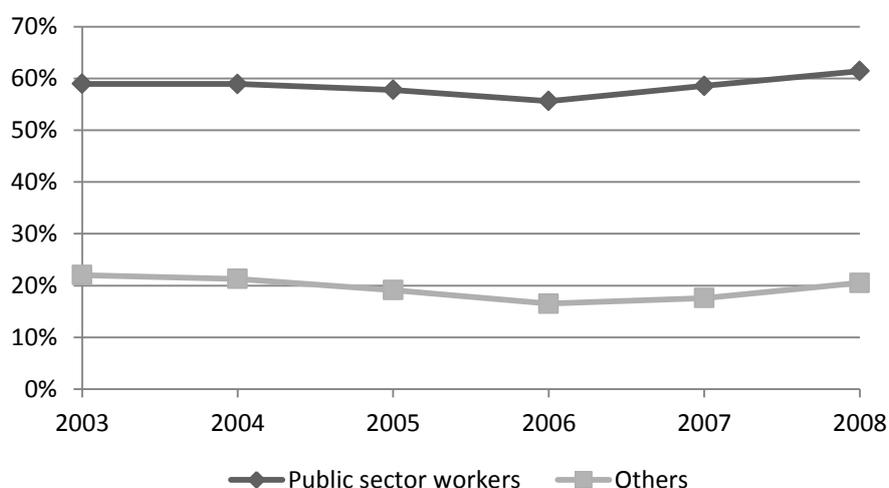
attenuation effects will cancel out to produce a consistent estimate of the causal effect of interest, even where some non-government workers are included in the CSS industry variable.

3.6 Results

Before performing the two-sample 2SLS, I plotted the trends in medical scheme coverage for which industry data is available in both the LFS/QLFS and GHS datasets. Figure 3.3, below, reveals that public sector employees are substantially more likely than other working aged South Africans to have medical aid in all periods, and that the pre-2006 trends are very similar. In fact, the hypothesis that these time trends are the same cannot be rejected ($p=0.1695$). The trends since 2006, on the other hand, appear to be diverging with a stronger increase in medical coverage amongst government employees. The hypothesis of a common post-GEMS trend is easily rejected (with p -value of <0.0001). I therefore proceeded to estimate equations [2] and [3] which apply the pre-GEMS common trend assumption.

Figure 3.3: Share of working age individuals with medical scheme coverage, 2003-2008

Source: LFS/QLFS, 2003-2008



The results from the combined analysis using the LFS/QLFS and the GHS are shown in Table 3.1 below. The GHS ceased to gather industry data for a number of years after 2008 and thus the analysis was limited to a two-year window shortly after the introduction of GEMS. There is a positive and significant coefficient on the 2008 year interaction with government sector, validating the IV approach. The F-test statistic is large, dispelling any concerns regarding weak instruments. According to the second stage estimates membership of medical schemes increases the likelihood that ill individuals will consult a health worker by 82%.

Table 3.1: Impact of insurance on utilisation, 2002 – 2008 (LFS/QLFS, GHS)

Variables	(1) Medical scheme	(2) Medical visits if ill
Public sector	0.219*** (0.00215)	-0.134* (0.0724)
Public sector*(Year = 2007)	0.0133*** (0.00483)	
Public sector*(Year = 2008)	0.0545*** (0.00365)	
Medical scheme (predicted)		0.824*** (0.312)
Observations	381,566	21,984
R-squared	0.272	0.025
Year effects	Y	Y
Fixed effects	N	N
Demographic controls	Y	Y
Education & Occupation	Y	Y
Geography	Y	Y
F-test	111.6	
Standard errors in parentheses		
*** p<0.01, ** p<0.05, * p<0.1		

Table 3.2 shows the first-step of the NIDS IV regression. The industry variable is based is self-reported and contains various contradictions. I therefore considered various options to filter the noise and compare the first-stage estimates across three options. The first column shows estimates for the raw variable prior to applying any filters. Government sector employees are 46% more likely to have health insurance than other (private sector or non-working) South Africans. The time coefficients demonstrate a small but significant decreasing trend in health insurance over time, which is not surprising given that the panel coincided with a recessionary period. The interaction between public sector employment and the wave 2 and 3 dummies shows that public sector employees were increasingly likely to obtain health insurance, although these effects are small and statistically significant. This result does not reflect the enrolment numbers from GEMS showing a steady steep trend (Figure 3.1).

Table 3.2: First-stage estimates of likelihood of being insured (NIDS 2008 – 2012)

Variables	(1) Medical scheme	(2) Medical scheme	(3) Medical scheme
Year 2010	-0.0223*** (0.00420)	-0.0223*** (0.00414)	-0.0170*** (0.00417)
Year 2012	-0.0309*** (0.00414)	-0.0320*** (0.00408)	-0.0204*** (0.00409)
Government	0.463*** (0.0110)		
Government 2010	0.0229 (0.0155)		
Government 2012	0.0116 (0.0152)		
Govt 2 years tenure		0.519*** (0.0118)	
Govt 2 years tenure 2010		0.0477*** (0.0168)	
Govt 2 years tenure 2010		0.0343** (0.0163)	
Government clean variable			0.519*** (0.0119)
Government clean var_2010			0.0408** (0.0175)
Government clean var_2012			0.0894*** (0.0185)
Constant	0.139*** (0.00283)	0.141*** (0.00279)	0.141*** (0.00282)
Observations	43,028	43,028	43,028
R-squared	0.119	0.133	0.116
F-test	1.095	4.339	11.70
Standard errors in parentheses			
*** p<0.01, ** p<0.05, * p<0.1			

There are at least two reasons why the estimates from the NIDS data could under-estimate the effect of GEMS on health insurance. Firstly, due to administrative bottlenecks and recall errors, there may be a lag between starting employment in the public sector and reporting that you have health insurance in a household survey. Secondly, and probably more seriously, if the public sector employment variable is captured with error then that will also downwardly bias the effect of GEMS in the estimates. I attempted to address these concerns by using alternative definitions of public sector employment. The regression reported in column 2 only uses public sector employees who reported having at least 2 years of tenure. In column 3, this variable is further cleaned using the reported employer from the other waves as well. For example, this definition ignores individuals who claimed to be unemployed in wave 1 and to have been employed in the public sector for 5 years in wave 2. The results from columns 2 and 3 show larger government-time interaction effects,

which confirms my concerns of health insurance and reporting lags and coding errors. The F-tests indicate that the instruments in column 3 have a strong relationship with the likelihood of having insurance and weak instruments are therefore not a concern.

Table 3.A1 (Appendix) investigates who benefitted most from GEMS in terms of an increased likelihood of having health insurance. The probability of being insured was estimated separately for government employees (column 1) and the rest of the population. Coverage is higher amongst females, the highly educated and members of high income households, and this pattern is more pronounced in the public sector than elsewhere. Coverage is also observed to be lower for blacks and coloureds than for whites and Indians, although these effects are weaker in the public sector. The interaction between time and these covariates for the public sector allow for an understanding of which groups experienced a sharper increase in insurance during the implementation of GEMS. Males and those with lower levels of education seem to have benefitted disproportionately. Although the public sector interactions with household income are not significant themselves, they are significantly different from the pattern for the rest of the population (where insurance amongst poorer household declined, potentially due to economic recession). This indicates that public sector employees in poor households experienced an increase in coverage relative to what would have happened in the absence of GEMS. In line with intuition and expectations, the effect of GEMS appears to have been larger for those groups who were the least likely to have been insured prior to its implementation.

The analysis starts with a pooled OLS regression on the determinants of having visited one of three types of healthcare facilities (private care, public care or no care) in the previous year. Due to endogeneity concerns, the coefficients reported in Table 3.3 are interpreted as partial correlations between equilibrium outcomes rather than treatment effects. Columns 1 and 2 consider the determinants of whether individuals opted to visit a private healthcare facility. Membership of a medical scheme is associated with a rise of 25% and 4% in the likelihood of consulting a private doctor and visiting a private hospital respectively. It is also associated with a 5% and 10% decline in the likelihood of visiting a public hospital and consulting a public doctor respectively. In the reported model specifications trends, demographic factors, education, income, reported symptoms and geography are controlled for.

The control variables have the expected signs. Being white or Indian, female, older, richer, better educated, reporting bad health, and living in urban areas or in the Free State, Western Cape, Mpumalanga or Gauteng provinces are associated with a higher likelihood of consulting private providers.

Using a comparable model specification, Table 3.4 reports the coefficients on the instrumented insurance variable from the second stage regression. These estimates show that belonging to a medical scheme increases the likelihood of seeking care (over the next 5 months) by 71%. It is encouraging that this estimate is comparable to that from the first IV strategy, which uses a different set of data (GHS and LFS) and considers an earlier phase in the life of GEMS (2007 and 2008).

Table 3.3: OLS estimates of effect of health insurance on choice of health facility (NIDS, 2008-2012)

Variables	(1) Private doctor	(2) Private hospital	(3) Public hospital	(4) Public clinic	(5) Other
Medical scheme	0.255*** (0.00508)	0.0372*** (0.00188)	-0.0480*** (0.00380)	-0.0945*** (0.00528)	0.000350 (0.000677)
Observations	49,980	49,980	49,980	49,980	49,980
Wave effects	Y	Y	Y	Y	Y
Fixed effects	N	N	N	N	N
Demographic controls	Y	Y	Y	Y	Y
Education & Income	Y	Y	Y	Y	Y
Symptoms	Y	Y	Y	Y	Y
Geography	Y	Y	Y	Y	Y

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 3.4 also shows a sizeable increase in the likelihood of consulting a private doctor (66%) and a private hospital (12%). There is a predicted increase in public hospital use and a strong decrease in public clinic use, but neither of these effects is significant.

The IV estimates are higher than the OLS estimates from Table 3.3. The OLS variables may be contaminated by measurement error and unobserved heterogeneity. Unobserved heterogeneity relating to adverse selection into insurance is expected to inflate OLS estimates relative to IV estimates. However, the IV estimates in Table 3.4 are larger than the OLS estimates, which could point towards a large role for measurement error. Alternatively, and perhaps more plausibly, the difference could be due to the IV capturing local average treatment effects (LATE) that attach more weight to those more affected by the GEMS policy. If this medical scheme drew new members mainly from households where individuals often chose to obtain no care and very rarely sought private care, then I could expect a larger effect on health seeking behaviour than for the OLS estimates, capturing marginal effects that are dominated by the traditional corps of insured individuals.

Table 3.4: 2SLS estimates of effect of health insurance on choice of facility (NIDS, 2008-2012)

Variables	(1) Any care	(2) Private doctor	(3) Private hospital	(4) Public hospital	(5) Public clinic	(6) Other
Medical scheme	0.705*** (0.207)	0.662*** (0.155)	0.121** (0.0550)	0.0602 (0.110)	-0.135 (0.151)	-0.00978 (0.0195)
Observations	49,990	49,980	49,980	49,980	49,980	49,980
Wave effects	Y	Y	Y	Y	Y	Y
Fixed effects	N	N	N	N	N	N
Demographic controls	Y	Y	Y	Y	Y	Y
Education & Income	Y	Y	Y	Y	Y	Y
Symptoms	Y	Y	Y	Y	Y	Y
Geography	Y	Y	Y	Y	Y	Y

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

3.7 Conclusion

The introduction of the GEMS in 2006 and gradual roll-out since 2007 created a natural experiment in the extension of health insurance eligibility. The aims of GEMS included improved equity and affordability of health insurance offered to government employees. It therefore involved an expansion of subsidy eligibility, a concerted effort to create more affordable benefit options for lower tier government workers and an increase of subsidies across the range, but also in particular for lower tier workers who qualified for a full subsidy for the most basic benefit option.

As expected in South Africa's polarised provider market, the extension of insurance caused a large response in both total utilisation and provider choice. The estimates from the two estimation strategies found similar results and provide evidence of a very large impact of insurance cover (71% to 82%) on the utilisation of health services. Insurance increases the likelihood of using private providers and in particular private doctors (66%).

These results are interpreted against the backdrop of a growing body of evidence showing strong preferences for private providers amongst South Africans. GEMS was launched partly to promote utilisation of government health facilities amongst government employees and to help generate additional revenue flows for public facilities. However, the research shows a strong shift away from government providers. Analysis of the scheme data shows that even amongst the low-salaried government employees, very few opted for basic benefit schemes that do not offer comprehensive private hospital cover.

While problems with the quality of public sector services are increasingly acknowledged, this analysis also clearly indicates the link between user acceptability (or even quality) and quantity. Providing healthcare services of greater acceptability levels to those without medical scheme coverage is likely to boost the per capita number of visits, which can yield significant public health benefits. For the “treated” sub-group, there appears to be few other significant constraints to demand that cannot be overcome if they can more acceptable healthcare services at a low cost. It is however important to bear in mind that while this sub-group may be distinct from the traditional core of medical scheme members, they are employed and therefore also distinct from the lowest quintiles of South African society. Poor households are likely to face harsher trade-offs and factors such as transport costs and childcare worries may be binding constraints for this group, even when the quality of healthcare services has vastly improved.

The analysis also offers useful inputs for the planning process supporting the pilot and launch of National Health Insurance. The large responses in utilisation and provider choice suggest that providing access to private doctors at no cost to the user is likely to cause a large increase in total utilisation, most of which will be directed towards these private doctors. Again, the demand response of the “treated” sub-group may overestimate the response of poor South Africans that face more constraints, but in lieu of other evidence such upper bound estimates can help set parameters for scenario-based forecasts to ensure adequate workforce planning. Given the pressure on public health budgets and the lack of doctors and nurses, these estimates also highlight the need for effective and fair rationing strategies and renewed emphasis on gatekeeping to accompany other NHI health reforms.

Limitations:

The main limitations of the analysis relate to the datasets used. However, the analysis demonstrates novel ways of overcoming these limitations. NIDS does not have a pre-GEMS implementation wave as NIDS only started in 2008 but this was overcome by supplementing the NIDS analysis with the two-sample estimation strategy using the LFS and GHS data. The QLFS does contain data on health seeking behaviour, necessitating a two-sample estimation strategy with health seeking data provided by the GHS survey. Furthermore, the identification strategy in both datasets relied on an imprecise variable to identify government sector workers. However, as argued earlier, any attenuation effects from this variable will cancel out over the 2SLS estimators to produce a consistent estimate of the causal effect of interest, even where some non-government workers are included in the CSS industry variable.

Implications for future research:

It is anticipated that the fourth wave of NIDS (NIDS 2014) will be released in the middle of 2016. This will allow the addition of a third wave to the analysis (relative to the baseline wave of NIDS 2008) to consider the impact of health insurance on the health seeking behaviour of government workers. Furthermore, it will also allow for consideration of the impact of health insurance on health outcomes. Health outcomes generally respond slowly to better access to health services and a fourth wave of the panel dataset may allow for the detection of impacts which are not present in the second or thirds waves.

Appendix A to Chapter 3

Table 3.A1: The effect of GEMS on health insurance, by sector

Variables	(1) Public sector	(2) Others
Log of personal income	0.175*** (0.0192)	0.0981*** (0.00222)
Female	0.0337 (0.0327)	0.000959 (0.00456)
Education	0.0249*** (0.00545)	0.00817*** (0.000622)
Black	-0.0892** (0.0445)	-0.306*** (0.00887)
Coloured	-0.121* (0.0677)	-0.288*** (0.0112)
Indian	-0.107 (0.0982)	-0.188*** (0.0166)
Year 2010	0.139*** (0.0449)	0.00889** (0.00345)
Year 2012	0.130** (0.0538)	-0.0218*** (0.00343)
Log of personal income 2010	-0.0112 (0.0272)	0.00309 (0.00332)
Log of personal income 2012	-0.00711 (0.0305)	0.0141*** (0.00333)
Female_2010	-0.110** (0.0495)	-0.00357 (0.00672)
Female_2012	-0.00909 (0.0529)	-0.00406 (0.00659)
Education_2010	-0.0179** (0.00831)	0.000398 (0.000924)
Education_2012	-0.0148* (0.00888)	0.000846 (0.000919)
Black_2010	-0.0361 (0.0703)	-0.0471*** (0.0144)
Black_2012	0.00712 (0.0788)	0.0482*** (0.0143)
Coloured_2010	-0.00119 (0.103)	-0.0612*** (0.0175)
Coloured_2012	0.177 (0.115)	0.0396** (0.0172)
Indian_2010	0.0963 (0.149)	0.0158 (0.0247)
Indian_2012	0.178 (0.157)	0.0702*** (0.0249)
Observations	1,825	41,132
R-squared	0.228	0.337

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Chapter 4

Gender and TB health seeking and detection: findings from a TB prevalence survey in the Western Cape

4.1 Introduction

In 2013 women accounted for 37% of all estimated new tuberculosis (TB) cases globally, with an estimated 3.3 million female incident⁴⁶ TB cases and 510,000 estimated female TB deaths (WHO, 2014). Although the female-to-male TB incidence ratio varies by country and region, globally about half as many women as men are notified with TB (WHO, 2014). Men may be more exposed to TB bacilli in their daily lives due to social and economic roles (Lienhardt, 2001) and may also be biologically more susceptible to developing TB (Neyrolles and Quintana-Murci, 2009). A large proportion of the difference in the female-to-male TB notification rate could also be due to the higher prevalence of smoking amongst men globally (Watkins and Plant, 2006). These points support the notion that the global female-to-male ratio of notified TB incident cases may be an accurate reflection of population incidence.

The debate on gender differences in TB case notifications, however, focuses not only on greater male exposure and biological susceptibility to TB, but also on systemic socio-economic and cultural gender biases that undermine the diagnosis and treatment of TB amongst women. Evidence indicates the health seeking behaviour (including provider choices) and diagnosis and care received by women with TB may be sub-optimal compared to the health seeking behaviour of and diagnosis and care received by men (Long et al., 1999; Wang et al., 2008; Karim et al., 2007; Eastwood and Hill, 2004; Miller et al., 2013; Begum et al., 2001).

South Africa has the third highest TB incidence in the world (WHO, 2014), with TB being the leading cause of reported natural deaths in South Africa (Statistics South Africa, 2014). It is estimated that 74% of the 2.4 million missed TB cases globally can be found in ten countries, with South Africa being one of the top five countries (WHO, 2014). South Africa also has the highest number of human immunodeficiency virus (HIV) positive TB cases in the world, with an estimated 270,000 incident cases in 2013 (WHO, 2014). In South Africa, 0.8 women are diagnosed with smear positive TB for every man diagnosed, compared to 0.5 women for every man globally (WHOa, 2013). The higher

⁴⁶ While in economic terms incidence refers to the total proportion of any population subject to a specific state, in health terms incidence or incident cases refers to new cases of disease detected over a defined period (in this case, a year).

rate of smear positive TB diagnosis amongst women in South Africa is most likely due to the disproportionate burden of HIV on women, particularly young women (Karim et al., 2009).

There is a global call for achieving gender equity through sex-specific health research at every level of research (Promoting equity through sex-specific medical research, 2014). I therefore examined gender patterns in the healthcare seeking behaviour and the TB diagnostic services using data from a TB prevalence survey in eight communities in the Western Cape, South Africa conducted as part of the Zambia, South Africa Tuberculosis and AIDS Reduction (ZAMSTAR) trial (Ayles et al., 2013).

While gender forms the main analytical focus of this paper it is, however, not the only important consideration. The detailed analysis of health seeking behaviour and TB diagnostic care received, relative to TB symptoms, whether through a gender lens or for all adults, allows for a more nuanced understanding of what the missed opportunities are in the detection of TB. The main and unique contribution of this chapter is to use data from a TB prevalence survey to examine health seeking behaviour through a gender lense. Data from TB prevalence surveys have rarely been used internationally for this purpose and, as far as I am aware, never in the South African context.

I start the chapter by reviewing literature on gender and health and gender and TB, respectively, and then move to provide more information on the study context and methods. Next, the study results are discussed and interpreted by referring to how the results relate to other study findings. I conclude by briefly referring to the policy implications of the findings.

4.2 Literature review⁴⁷

In this section I briefly consider why the issue of gender deserves attention in a health and development context, while also summarising some of the most prominent evidence on gender differences and patterns in TB health seeking, as well as some of the factors that may influence these differences. In considering gender differences in health and health seeking (also in a TB context), I emphasise the behaviour of women (relative to those of men). *Ex ante* there is no reason, however, to believe that women will necessarily be at a disadvantage relative to men.

⁴⁷ This section is informed by a literature search which used a variety of electronic search engines, including Google, Google Scholar, Econlit and ScienceDirect. Some of the most prominent journals in health economics (Journal of Health Economics and Health Economics) were also searched directly. The main search terms that were used include “gender”, “sex”, “health”, “disease” and “tuberculosis” and “TB” in a variety of combinations. Once initial important sources were found, their bibliographies were also consulted for additional references.

4.2.1 Why gender matters for health

In considering the role of gender in health, it is important to distinguish between biological sex and gender (Sen, Östlin, and George, 2007; Buvinić et al., 2007). While biological sex is determined on chromosome level, gender refers to the “social power relations and hierarchies” that are influenced by factors such as “biological sex, age, life-cycle position, and family status” (Sen, Östlin, and George, 2007: 7). This is of importance in a health context, because gender, like socio-economic status, can be a source of inequity in health outcomes. Equality (or “sameness”) in health outcomes between men and women, however, does not point towards the absence of health inequity due to gender. Ensuring gender equity in health typically require the following approach (Sen, Östlin, and George, 2007:7, citing the work of Breen, 2002, Iyer et al., 2007, Sen et al., 2002):

“Where biological sex differences interact with social determinants to define different needs for women and men in health (the most obvious being maternity), gender equity will require different treatment of women and men that is sensitive to these needs. On the other hand, where no plausible biological reason exists for different health outcomes, social discrimination should be considered a prime suspect for different and inequitable health outcomes. Health equity in the latter case will require policies that encourage equal outcomes, including differential treatment to overcome historical discrimination.”

Gender, as defined above, can be a source of inequity in both exposure to and the treatment and detection of disease. It is important to distinguish between biological susceptibility due to biological sex and vulnerability due to gender (Sen, Östlin, and George, 2007), the latter being a state influenced by socio-economic correlates and power structures in society.

In this context, the achievement of equality of opportunity is required. Equality of opportunity can best be understood as “the assignment of individuals to places in the social hierarchy...[being]... determined by some form of competitive process, and all members of society are eligible to compete on equal terms” (Arneson, 2015). The health system and its engagement with men and women should therefore be structured in such a way as to provide both genders with equal opportunity to utilise health services, but still allow for differences in individual behaviour and risk preferences as long as those preferences are not systematically influenced by gender.

Women’s vulnerability in health choices and options becomes clearer in the context of healthcare seeking behaviour. Before seeking care for any symptoms or disease, women first need to be aware that they are experiencing a health problem (Sen, Östlin, and George, 2007). Awareness of certain health problems may be low due to low health knowledge and also general education levels

compared to those of men and this plays a role in healthcare seeking and utilisation (O'Connell et al., 2015). Once the existence of a health problem has been identified, the problem needs to be acknowledged (Sen, Östlin, and George, 2007). However, women more so than men, may suffer from stigma and marginalisation for certain diseases and purposefully elect not to seek treatment as this will be an acknowledgement of the problem (Sen, Östlin, and George, 2007; Vlassoff, 2007).

When women finally decide to seek care for their symptoms or problem, they may experience greater access barriers, e.g. affordability and availability (physical distances), than men in accessing healthcare (Buvinić et al., 2006; Sen, Östlin, and George, 2007; Vlassoff, 2007). It is possible that they may also make sub-optimal healthcare provider choices, such as self-treatment or consulting alternative healers for their specific healthcare problem due to affordability or availability barriers (Vlassoff, 2007). Women also experience abuse by healthcare workers, especially in seeking care for reproductive health, which may influence their decision to consult with a public rather than private or alternative healthcare provider (Målqvist et al., 2012).

Women's ability to seek care for different health conditions may be restricted by gender norms that influence autonomy to seek care (Hawkes and Buse, 2013). In fact, decision-making and autonomy as related to gender has been identified as one of five main categories of non-financial access barriers to healthcare⁴⁸ in an international review of qualitative and quantitative literature on access barriers (O'Connell et al., 2015). Autonomy is not isolated from broader socio-economic and cultural factors, of which an important determinant is gender differences in the allocation of time use and household responsibilities. The largest share of the childcare and household management burden is often allocated to women and this can potentially have quite a large influence on women's healthcare seeking behaviour. Young (1996: 1959) finds "the double burden of caring and paid work makes it difficult for all women, not just those on a low incomes, to fit primary and preventative health care into their daily activity schedule" amongst a sample of Liverpool women. A delay in seeking healthcare for possible cancer symptoms has been found to be associated with childcare responsibilities (Facione and Facione, 2006), while an association has been found between HIV-positive women living with more children (younger than 18) in the household and lower ARV adherence rates (Merenstein et al., 2008; Merenstein et al., 2009).

It is, however, important to keep in mind the potential role of confounding factors in considering the impact of gender on health seeking behaviour. Health or epidemiological studies rarely employ

⁴⁸ The other categories include ethnicity, religion, physical accessibility and education, knowledge and information.

causal estimation methods and many of the findings that inform the above discussion is based on bivariate or, at most, simple multivariate analysis.

4.2.2 Gender and TB

After reviewing five quantitative studies⁴⁹ on TB and health care seeking behaviour, Van den Hof et al. (2010) conclude that women, in contrast to men, more often seek access to healthcare when they experience possible TB-related symptoms. This conclusion seems to also hold in South Africa, at least also in urban South Africa. In a cross-sectional study of 104 TB patients, Meintjes et al. (2008) found that there is a significant association between delayed access to healthcare and being male.

Provider choices:

Apart from potentially seeking healthcare earlier than men, women may differ from men in terms of the types of providers they tend to seek care from. Some studies find that men are more likely to seek (actual behaviour) or intend to seek care from higher levels of care such as hospitals (rather than public clinics) in The Gambia (Lienhardt et al., 2001), China (Wang et al., 2008) and Vietnam (Hoa, Chuc and Thorson, 2009). In the study by Hoa, Chuc and Thorson (2009) in Vietnam where intended care seeking was reported by a sample of women and men, significantly more men reported they would first seek care at the district hospital. However, there was also a significant association between average TB knowledge and intended hospital care seeking, with men found to have significantly higher knowledge levels of TB (Hoa et al., 2009).

Women have been found to be more likely to consult providers that operate on lower levels of care, which may lead to longer delays in getting appropriate access to healthcare. An analysis of adults (n=492) who had coughed for three weeks or more amongst a larger sample of adults in Vietnam (Thorson Hoa and Long, 2000) found a significantly larger number of health seeking actions (for cough symptoms) amongst women. The study also found that women were significantly more likely to seek care from lower level or less qualified providers for TB diagnosis such as self-treatment, a pharmacist or a private practitioner⁵⁰. The reasons for provider choice more commonly reported by women than men in this study include “convenience and close proximity to home” (Thorson, Hoa and Long, 2000: 1823). A study amongst a small sample (N=152) of adults diagnosed with TB in The Gambia also found a significantly higher number of health seeking actions amongst women compared to men (Lienhardt et al., 2009). In this context, however, although men were more likely to consult at hospitals than women, men were also more likely to consult with a pharmacist or

⁴⁹ The studies were conducted in Vietnam, China and the United States (amongst Mexican immigrants).

⁵⁰ Public health facilities are best positioned to diagnose TB given the public health nature of the disease.

doctor. Women were significantly more likely to consult at public health centres (Lienhardt et al., 2009). In a sample of adults in rural China who coughed 3 weeks or longer, women were also significantly more likely to seek care from lower level providers such as rural village clinics and pharmacies (Wang et al., 2008).

Factors influencing healthcare seeking decisions:

At least four categories of factors seem to have some type of influence on women eventually receiving appropriate healthcare when infected with TB: TB knowledge and awareness; stigma or perceptions of stigma associated with TB; access to or control over resources; and different household roles and responsibilities to men. More information on these factors are provided below.

In a cross-sectional survey (N=12,143) of adults in Vietnam (Hoa, Chuc and Thorson, 2009) a significant association was found between gender and TB knowledge levels. In multiple linear regression analysis which controlled for general education levels, age, socio-economic status and occupation using this data, men were significantly more likely to have higher TB knowledge levels than women. A cross-sectional survey in China (N=1,083) also found significantly higher levels of TB knowledge amongst men – this included questions on knowledge of a cough of 3 or more weeks as the main TB symptom, the local facility where TB would be treated, that TB would be treated free of charge and that TB could be cured (Wang et al., 2008). This study did not, however, control for differences in average education levels between men and women.

Women potentially face greater stigma or have perceptions of greater stigma due to a TB diagnosis. In a large cross-sectional survey of Vietnamese adults, Hoa, Chuc and Thorson (2009) found women were more likely to report associating stigma with TB, while a significantly larger percentage of women reported that they would hide a TB diagnosis from people in their support structures. Findings from qualitative research also confirm the possibility of greater stigma for women. In a number of focus group discussions in Vietnam, women more so than men (Long et al., 2001) voiced concerns about the social consequences such as isolation from family members and even divorce due to a possible TB diagnosis. In qualitative semi-structured interviews with both health workers and male and female patients in The Gambia, health workers reported cases of women being divorced by their husbands due to a TB diagnosis, while female TB patients, more so than male patients, spoke about stigmatisation experiences due to their disease (Eastwood and Hill, 2004).

In relative terms, travel and treatment could come at a greater cost to women due to lower incomes and limited control over household income. In an analysis of adults who coughed ≥ 3 weeks in a larger population-based sample of adults in Vietnam (Thorson, Hoa and Long, 2000), it was found

that in cases where adults did seek care, women reported much lower levels of healthcare expenditures for their selected healthcare providers. Given that the study also found that women typically consulted with less qualified providers or made sub-optimal care seeking choices (e.g. self-treatment), this could be indicative of the larger relative cost of TB care seeking for women. In qualitative semi-structured interviews with health workers in The Gambia, health workers reported transport fees to healthcare facilities as a greater access burden for women than men in receiving TB treatment (Eastwood and Hill, 2004).

Lastly, delay in seeking care for possible TB symptoms has also been ascribed to women's limited time and greater household responsibilities, including looking after children in various cultural settings (Johansson, 2000; Karim et al., 2007; Weiss et al., 2006).

Symptoms and sputum submission for testing:

Women may differ from men also in terms of the TB symptoms with which they present, making TB detection more difficult given healthcare workers' understanding of the typical TB patient's symptoms. In a sample (N=173) of individuals diagnosed with TB in the United Kingdom, Breen et al. (2008) found in multivariate analysis that women were significantly less likely to report systemic symptoms typically associated with TB, including fever, weight loss and night sweats. In a sample of adults (aged 15 years and older, N=1,027) of diagnosed TB patients in Vietnam, Long Diwan and Vinkvist (2002) found that women were significantly more likely not to report coughing, the production of sputum expectorations and the coughing up of blood.

Studies have also found that women tend to produce poorer quality sputum samples. It is hypothesised that this may be due to their more genteel (and possibly gendered) approach to coughing. In a Kenyan study amongst a small sample of potential TB patients (N=644), significantly fewer women than men were able to produce a set of three sputum samples considered "good quality", influencing the ability of the system given the guidelines in place then to detect TB amongst these sputum samples (Ramsay et al., 2009). When detailed sputum submission instructions were provided to women (N=1,494) in Pakistan, the TB case detection rate using smear microscopy for this group increased significantly (Khan et al., 2007). In this study, more detailed sputum submission instructions provided to men (N=1,561) was not associated with a significantly increased TB detection rate amongst men.

4.3 Study context

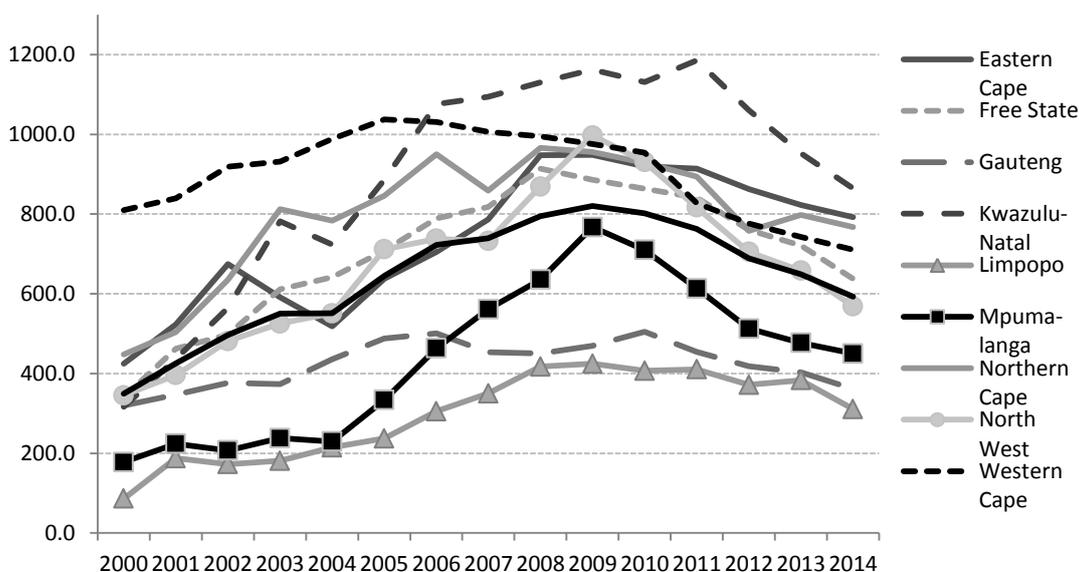
The data analysis (also see Section 4.4 on data) is based on a TB population prevalence survey of adults (aged ≥ 18 years) in eight high TB prevalent communities in the Western Cape, South Africa. I

briefly provide some information on the nature of the TB burden and cure rate in the Western Cape compared to other provinces. The gender distribution is also considered.

While the Western Cape was the province with the largest population-level TB incidence between 2000 and 2005, by 2006 Kwazulu-Natal took the position of the province with the highest population-level TB incidence (Figure 4.1). In 2010, the ZAMSTAR prevalence survey year, the Western Cape had the second highest TB incidence rate in South Africa. By 2014, however, the population-level TB incidence in the Western Cape had decreased sufficiently for it to only have the fourth largest TB incidence burden.

Figure 4.1: Reported cases of all TB (pulmonary, extra-pulmonary, drug-resistant, extremely drug-resistant) per 100,00 of the population as obtained from TB registers

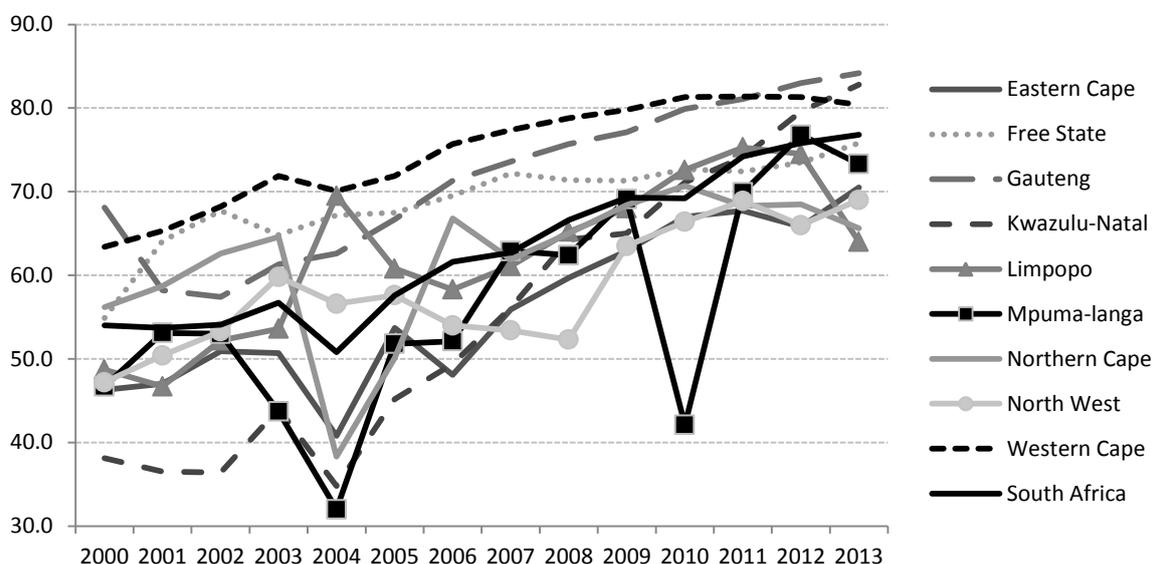
Source: DHIS, Department of Health as reported by Health Systems Trust



The TB cure rate can be considered a rough proxy of health system effectiveness in dealing with diagnosed TB. In 2010, the ZAMSTAR prevalence survey year, the Western Cape, followed by Gauteng and then Limpopo, had the highest TB cure rate of all the provinces (Figure 4.2). By 2012, its performance in terms of the TB cure rate had, however, been surpassed by Gauteng and by 2013 both Gauteng and Kwazulu-Natal had higher TB cure rates than the Western Cape.

Figure 4.2: TB cure rate for new smear-positive cases as obtained from TB registers

Source: DHIS, Department of Health as reported by Health Systems Trust



While gender breakdowns of various categories of TB diagnoses are not reported into the District Health Information System (DHIS), the data capturing system of South Africa's public health sector, this can be obtained from household surveys. The GHS asks respondents who indicated they were ill or injured in the month preceding the survey about their illness or injury. Respondents were able to indicate that they had "TB or severe cough with blood" as one of 15 answer options. Due to stigma, survey respondents may, however, be reluctant to report having TB. The nature and position of the question could also only elicit responses from individuals who had TB at a more advanced or severe stage of the disease.

Pooling data from the GHS for several years could, however, provide a sufficiently large sample to explore at least some of the socio-demographic characteristics of individuals who report having TB, even though not fully representative of all people with TB in South Africa at any given moment. I pooled data from 2002 to 2011, which provided 6,568 (unweighted) individuals or observations who reported having TB. This equates to 0.7% of the total of 1,000,425 individuals in the dataset for the period. Given the small sample size of individuals who report TB over the total period, it was not possible to conduct bivariate analysis for specific years, but bivariate analysis for one or at most two socio-demographic characteristics of individuals who reported TB was done. Eight percent (Table 4.1) of the sample who reported having TB reside in the Western Cape.

Table 4.1 Provincial distribution of TB sample*Source:* Pooled GHSs, 2002-2011

Province	%	95% Confidence Interval
Western Cape	8.0	7.2-9.0
Eastern Cape	19.1	17.9-20.3
Northern Cape	2.3	2.1-2.6
Free State	7.0	6.3-7.7
Kwazulu-Natal	27.5	26.2-28.9
North West	10.1	9.3-10.9
Gauteng	12.8	11.6-14.0
Mpumalanga	6.6	6.0-7.2
Limpopo	6.7	5.8-7.7

Of individuals who reside in the Western Cape and who have report having TB in the survey, 52.8% (Table 4.2) are men. However, given overlapping confidence intervals with the percentage of women in the Western Cape who report having TB, it cannot be concluded that there are significantly more men in the Western Cape than women who report having TB at population-level. The only provinces which do not have overlapping confidence intervals and therefore statistically significant differences in reported TB between men and women are the Eastern Cape, Kwazulu-Natal, Mpumalanga and Limpopo.

Table 4.2 Gender distribution by province of adults that reported being ill with TB*Source:* Pooled GHSs, 2002-2011

Province	Men (%)	Men 95% Confidence Interval	Women (%)	Women 95% Confidence Interval
Western Cape	52.8	46.9-58.7	47.2	41.3-53.2
Eastern Cape	53.5	41.3-53.2	46.5	43.2-49.9
Northern Cape	59.6	53.8-65.0	40.4	35.0-46.2
Free State	57.2	52.2-62.0	42.8	38.0-47.8
Kwazulu-Natal	52.2	49.3-55.1	47.8	45.0-50.8
North West	55.2	51.1-59.3	44.8	40.8-48.9
Gauteng	59.3	54.2-64.3	40.7	35.8-45.8
Mpumalanga	48.4	43.8-53.1	51.6	46.9-56.2
Limpopo	47.9	40.6-55.3	52.1	44.7-59.4

4.4 Data

The TB prevalence survey for the ZAMSTAR trial was conducted between January 9, 2010 and December 6, 2010 as the primary endpoint measuring the impact of the ZAMSTAR trial interventions implemented between August 1, 2006, and July 31, 2009 (Ayles et al., 2013). The trial was focused on enhancing TB case detection outside primary healthcare facilities through community interventions and had four arms. All participating Western Cape communities in the trial had an enhanced TB-HIV programme implemented at participating primary healthcare facilities in their areas. The first trial arm had no further intervention and served as a control group. The second arm had enhanced community TB case finding, while the third arm had specific household interventions. The fourth arm had a combination of enhanced community TB case finding and household interventions (Ayles et al., 2013).

Enumerator areas in the eight communities were randomly selected. Field workers visited households to obtain head of household permission for survey participation. Once permission was provided, all adults were enumerated and individual written informed consent obtained from adults for enrolment in the survey.

Data on socio-economic and demographic variables, TB symptoms, current and previous TB treatment were collected. In line with the 2009 National Tuberculosis Management Guidelines of South Africa, a cough of ≥ 2 weeks was regarded as a symptom of possible TB (Department of Health, 2009). Health seeking behaviour questions included questions on the current presence of a cough, the duration of the cough and any actions taken, including provider choices, to seek help for the cough. Once participants indicated that they had sought help, they were asked whether the healthcare provider collected a sputum sample, whether they provided one and about the outcome of the sputum test. Participants were also asked whether they were currently on TB treatment.

Sputum samples were obtained from all adults (with and without symptoms) and inoculated onto manual mycobacterial growth indicator tubes (MGIT, Becton Dickinson and Company, New Jersey). Participants were considered to have a culture positive for *Mycobacterium tuberculosis* (*M.tb*) if the MGIT culture was positive and confirmed to be *M.tb* by 16SrRNA sequencing. Participants were encouraged to test for HIV using a finger-prick whole blood collection and rapid HIV testing (Determine as screening test and Unigold as confirmatory test) according to the standard of care at the time, either in the household or in mobile voluntary counselling and testing (VCT) centres in the community.

Further detail on the methodology and the larger ZAMSTAR trial has been published elsewhere (Ayles et al., 2013).

If loss of adults from the enumerated sample due to only non-consent is taken into account, the survey had a response rate of 77.7%. However, if loss of non-consenting adults, questionnaires with too many missing questions and non-evaluable sputum samples are taken into account, the survey had a response rate of 71.1%. Observations were available for a total of 30,017 adults.

Ethics:

The ZAMSTAR trial, including the prevalence survey, was approved by the ethics committees of Stellenbosch University (N04/10/173), the London School of Hygiene and Tropical Medicine, and the University of Zambia. The Western Cape Provincial Department of Health provided permission for the trial to be conducted in health facilities and surrounding communities. Participants in the prevalence survey provided written informed consent for TB testing, with separate written informed consent for HIV testing.

4.5 Methods

As I was mainly interested in establishing gender differences, the analysis focuses on the differences between gender and any other variable, e.g. gender and TB prevalence. Vaus (2002) recommends using a statistical method designed for bivariate analysis when considering the difference between any two variables. According to Aschengrau and Seage (2014) the chi-square test (χ^2) is commonly used to establish the statistical significance of the difference between two variables. The mean population values for men, women and the total population, as well as p-values for gender differences for a variety of variables were therefore established through the the chi-square test (χ^2). P-values <0.05 were considered to be statistically significant.

A step-wise approach was used to analyse the self-reported health seeking behaviour of the total sample (all adults) and this approach was replicated in gender-based analyses for, respectively, all adults who reported a cough of ≥ 2 weeks and culture-positive adults who reported a cough of ≥ 2 weeks.

I distinguished between symptomatic and non-symptomatic adults, identified how many of the symptomatic adults sought care and, for those who sought care, which provider they consulted. For adults who reported consulting at a primary healthcare facility prior to the prevalence survey, it was identified whether they were offered a sputum test, provided a sputum sample and what the reported result was.

4.6 Results

A total of 30,017 adults completed the questionnaire, had evaluable sputum samples and were included in the analysis.

Symptomatic survey participants:

Of these 30,017 adults, 1,670 (5.4%) coughed ≥ 2 weeks (Table 4.3). Five hundred and four (30.2%) said they consulted someone about their cough, of which 403 (80%) reported having visited a primary healthcare facility, 9 (1.8%) consulted a district hospital and 4 (0.8%) consulted with staff at the ZAMSTAR trial intervention sputum collection point during the intervention period, before the prevalence survey (Table 4.4). Eighty eight (17.4%) used a private healthcare service of whom 78 (15.4%) consulted a private clinic or doctor and 10 (2%) a pharmacy (Table 4.4). There were no significant differences by gender in terms of the type of healthcare provider selected (Table 4.4).

Table 4.3: Health seeking behaviour and TB diagnostic care received by adults who coughed ≥ 2 weeks (by gender)

	Men (N=11,297) n(%)	Women (N=18,720) n(%)	All (N=30,017) n(%)	P-value
Coughed ≥ 2 weeks	720 (6.4)	950 (5.1)	1,670 (5.6)	<0.001
Consulted someone	194 (26.9)	310 (32.6)	504 (30.2)	0.012
Went to primary healthcare facility	158 (81.4)	245 (79.0)	403 (80.0)	0.511
Was asked for sputum sample	122 (77.2)	155 (63.3)	277 (68.7)	0.003
Provided sputum sample	116 (95.1)	151 (97.4)	267 (96.4)	0.301
Positive sputum result	24 (20.7)	19 (12.6)	43 (16.1)	0.023
On TB treatment at time of survey	11 (45.8)	13 (68.4)	25 (55.8)	0.139

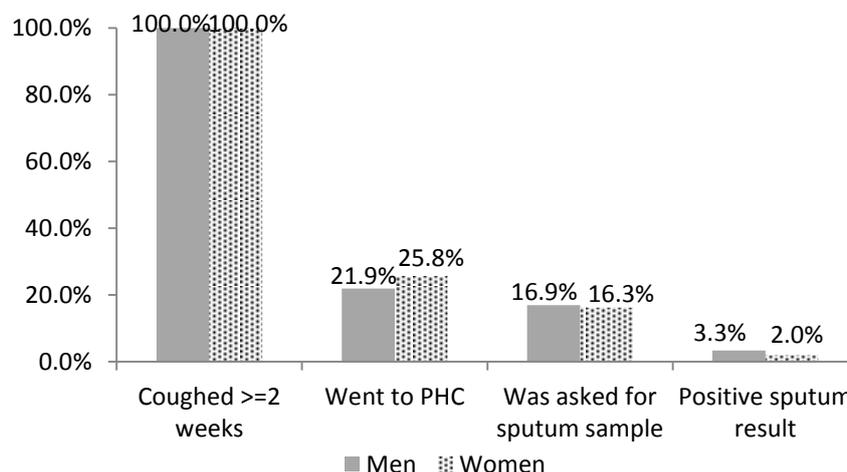
Table 4.4: Provider choice by gender for adults who coughed ≥ 2 weeks and sought care

	Men (n=194) n(%)	Women (n=310) n(%)	All (n=504) n(%)	P-value
Primary healthcare facility	158 (81.4)	245 (79.0)	403 (80.0)	0.511
District hospital	5 (2.6)	4 (1.3)	9 (1.8)	0.288
Private clinic or doctor	27 (13.9)	51 (16.5)	78 (15.5)	0.444
Pharmacy	2 (1.0)	8 (2.6)	10 (2.0)	0.225
ZAMSTAR sputum collection point	2 (1.0)	2 (0.6)	4 (0.8)	0.635

Men were significantly more likely ($P < 0.001$) to report coughing for ≥ 2 weeks: 720 men (6.4%) compared to 950 women (5.1%) (Table 4.3). Of the symptomatic men, 194 (26.9%) sought care compared to 310 (32.6%) of the symptomatic women ($P = 0.012$), while 158 (21.9%) of symptomatic

men and 245 (25.8%) of symptomatic women sought care at a primary healthcare facility (Table 4.3 and Figure 4.3).

Figure 4.3: Health seeking and TB diagnostic cascade of adults who coughed ≥ 2 weeks (by gender; 11,297 men and 18,720 women)*



*Percentages in figure calculated relative to total adults who coughed ≥ 2 weeks

Hundred and twenty two (77.2%) of the symptomatic men who reported visiting a primary healthcare facility were asked to submit a sputum sample compared to 155 (63.3%) women ($P=0.003$) (Table 4.3). A total of 24 men, 20.7% of those tested, reported a positive sputum smear result compared to 19 (12.6%) women of those tested ($P=0.023$). This equates to 2% of symptomatic women and 3.3% of symptomatic men being diagnosed with smear positive TB by primary healthcare facilities (Figure 4.3). Eleven (45.8%) symptomatic men and 13 (68.4%) symptomatic women who reported testing smear positive also indicated receiving TB treatment at the time of the survey ($P=0.139$) (Table 4.2). It is not clear whether they were put on TB treatment as a direct result of their reported interaction with the health system.

There were no significant differences between symptomatic men and women who sought healthcare in terms of age, highest level of education or occupation (Table 4.5). A significantly larger proportion of symptomatic men (49.4%) compared to women (33.9%) indicated they had previously been treated for TB ($P=0.002$). Symptomatic men were significantly more likely to mention symptoms of fever, weight loss and night sweats ($P=0.001$, $P=0.055$ and $P=0.001$).

Table 4.5: Characteristics of adults (by gender) in prevalence survey who coughed ≥ 2 weeks and sought care at a primary healthcare facility

	Men (n= 158) n(%)	Women (n=245) n(%)	P-value
Age category (years):			0.842
18-24	22 (13.9)	28 (11.4)	
25-34	32 (20.3)	60 (24.5)	
35-44	42 (26.0)	62 (25.3)	
45-54	35 (22.2)	50 (20.4)	
55+	28 (17.7)	45 (18.4)	
Highest educational attainment			0.672
No education	13 (8.2)	23 (9.3)	
Some primary school	54 (34.2)	71 (29.0)	
Some high school	70 (44.3)	123 (49.0)	
High school completed	18 (11.4)	29 (11.8)	
Tertiary education	3 (1.9)	2 (0.8)	
Main occupation			0.086
Employed	48 (30.4)	63 (25.7)	
Casually employed	18 (11.4)	22 (9.0)	
Unemployed or own land	76 (48.1)	127 (51.8)	
Student or homemaker	9 (5.7)	23 (9.4)	
Unable to work	7 (4.4)	10 (4.1)	
Previously treated for TB	78 (49.4)	83 (33.9)	0.002
HIV status known	106 (67.1)	156 (63.7)	0.483
Other symptoms:			
Chest pain	87 (55.1)	128 (52.2)	0.580
Fever	99 (62.7)	111 (45.3)	0.001
Weight loss	96 (60.8)	125 (51.0)	0.055
Shortness of breath	75 (47.5)	107 (43.7)	0.455
Night sweats	99 (62.7)	111 (45.3)	0.001

Culture-positive survey participants:

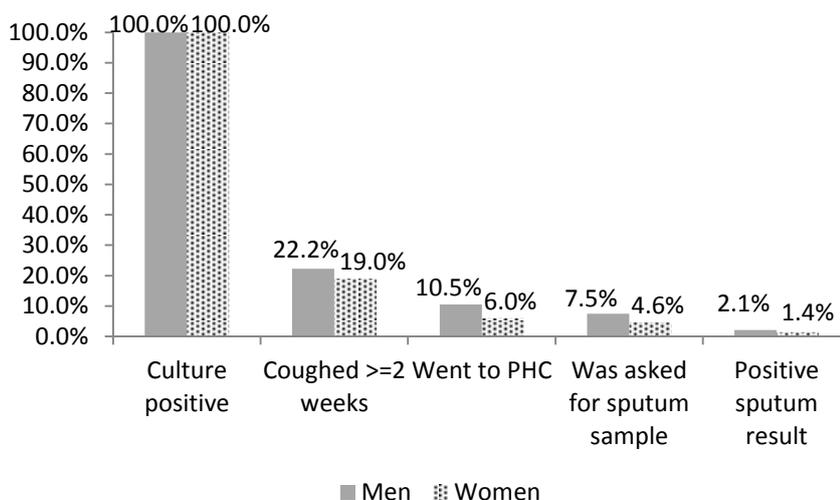
Of the 30,017 participants in the survey, 702 had culture-positive TB (Table 4.6). Men were significantly more likely to have culture-positive TB - 333 of 11,297 men (3%) were culture positive compared to 369 (2.0%) women ($P < 0.001$). Seventy four (22.2%) symptomatic men and 70 (19.0%) symptomatic women with culture-positive TB reported coughing for ≥ 2 weeks (Table 4.6 and Figure 4.4, $P = 0.474$). Culture-positive women were marginally less likely than men to report seeking any type of care for their cough ($P = 0.086$). Twenty six (37.1%) women indicated they sought care for their cough compared to 38 (51.4%) men (Table 4.6).

Table 4.6: Health seeking behaviour and TB diagnostic care received by culture-positive TB adults in prevalence survey (by gender)

	Men (N=11,297) n(%)	Women (N=18,720) n (%)	All (N=30,017)	P-value
Culture-positive TB cases	333 (3.0)	369 (2.0)	702 (2.3)	<0.001
of which:				
Coughed for two weeks or more	74 (22.2)	70 (19.0)	144 (20.5)	0.474
Consulted someone	38 (51.4)	26 (37.1)	64 (44.4)	0.086
Went to primary healthcare facility	35 (92.1)	22 (84.6)	57 (89.1)	0.346
Asked for sputum sample	25 (71.4)	17 (77.3)	42 (73.7)	0.626
Gave sputum sample	24 (96.0)	17 (100.0)	41 (97.6)	0.404
Positive sputum result	7 (36.8)	5 (33.3)	12 (35.3)	0.832
On TB treatment at time of survey	6 (85.7)	4 (80)	10 (83.3)	0.793

Of those with culture confirmed TB who reported a cough of ≥ 2 weeks and sought care, 35 men (92.1%) and 22 women (84.6%) consulted at a primary healthcare facility (Table 4.6). This equates to 10.5% of men and 6% of women with culture-positive TB seeking care for their cough symptoms at a primary healthcare facility (Figure 4.4). Once they consulted at a primary healthcare facility, 25 men (71.4%) and 17 women (77.3%) were asked for a sputum sample. Twenty four men (96%) and 17 women (100%) of those asked for a sputum sample provided one. Self-reported data on the outcome of sputum tests indicate that primary healthcare facilities diagnosed 7 (2.1%) men and 5 (1.4%) women with culture-positive TB, as smear positive TB cases (Figure 4.4).

Figure 4.4: Health seeking and TB diagnostic cascade of culture-positive adults (by gender; 333 men and 369 women)*



*Percentages in figure calculated relative to total culture-positive adults

Given the small sample size of culture-positive adults who sought care at a primary healthcare facility, it was difficult to determine the significance of demographic and socio-economic factors, as well as the significance of the self-reported presence of other symptoms. I therefore do not report it here.

The same step-wise gender-based analysis was repeated for sub-samples of symptomatic adults who were HIV-positive and adults who reported having previously had TB. No evidence was found of HIV status or reported previous TB driving the reported patterns.

4.7 Discussion

The study found that only a small percentage (30.2%) of adults who reported coughing for ≥ 2 weeks sought any type of care, with a higher percentage (44.4%) of adults with culture-positive TB seeking care for their cough. Furthermore, slightly more than two thirds (68.7%) of adults who reported seeking care for their cough at a primary healthcare facility indicated they were offered a sputum test at these facilities, despite current recommendations that all such patients should be offered a sputum test.

While women were significantly more likely to seek healthcare for their cough, they were significantly less likely to be asked for a sputum sample when they presented at a primary healthcare facility. They were also significantly less likely to report a positive sputum result.

The finding on the health seeking behaviour of women who reported coughing for ≥ 2 weeks is similar to other studies which found women are more likely than men to take healthcare action when experiencing TB symptoms (Hoa et al., 2009; Thorson, Hoa and Long, 2000; Wang et al., 2008). The implied lower likelihood of men with a cough of ≥ 2 weeks to seek care is also consistent with an earlier study in the Western Cape that found patient delay in seeking care for TB is greater amongst men (Meintjes et al., 2008).

Studies in India, Vietnam and The Gambia found that women prefer consulting with traditional healers, pharmacies and private providers given their greater convenience, proximity and privacy (Eastwood and Hill, 2004; Hoa et al., 2009; Thorson, Hoa and Long, 2000; Weiss et al., 2008). I, however, found no differences between men and women with regards to the type of healthcare provider selected.

In contrast to the health seeking behaviour of all women who reported cough symptoms, culture-positive women who experienced similar symptoms were marginally less likely than men to seek care. The data did not allow me to fully explore the socio-economic characteristics of the latter

group of women, but it may be that they are fundamentally different from the larger group of women. This requires further investigation.

TB cases amongst adults who present at facilities for reasons other than TB or who accompany children or other adults to these facilities are going undetected (Claassens et al., 2013). This analysis demonstrates that TB cases are also going undetected amongst adults who specifically seek care for possible TB symptoms at primary healthcare facilities. The SA National TB Programme Guidelines of 2009 identified primary healthcare facilities as the main entry point to TB diagnosis and care in South Africa (Department of Health, 2009). Weak implementation of screening protocols can undermine the identification of TB cases and contribute to the high number of missed TB cases in South Africa.

However, the data presented here also shows that a large majority of culture-positive TB identified through the community-based prevalence survey of the ZAMSTAR trial did not cough for at least two weeks and even more did not seek TB diagnosis and care through primary healthcare facilities. It is therefore important to continue exploring the value of community-based interventions seeking to identify TB outside primary healthcare settings.

The prevalence survey does not allow for the identification of the causes of weak or inconsistent application of screening protocols to patients of different genders. A Ugandan study reports similar gender differences in requests by health workers for sputum samples and speculates that this may be due to stereotypes held by health workers of the typical TB patient being male rather than female (Miller et al., 2013). Ultimately, further investigation of health worker perception and biases with regards to the gender of the typical TB patient will be required.

In addition to women having a lower likelihood of being asked for a sputum sample, women were also less likely to report having had a positive test result. The latter finding confirms an earlier study that analysed laboratory data on sputum tests in the Western Cape (Austin et al., 2004). It is necessary to establish whether this outcome is due to the provision of sub-standard sputum samples by women or to actual differences in TB incidence rates amongst men and women. Women have been found to be significantly more likely to deliver poor quality sputum samples due to their more genteel approach to coughing and sputum production (Khan et al., 2007; Ramsay et al., 2009). When detailed sputum submission instructions were provided to women in Pakistan, the TB case detection rate using smear microscopy for this group increased significantly (Khan et al., 2007).

Women who presented at a primary healthcare facility after coughing for ≥ 2 weeks were less likely to report other typical TB symptoms, including fever, weight loss and night sweats and more men

reported positive TB test results from sputum examinations. A greater presence of typical TB symptoms amongst men would thus be consistent with more men being TB positive or being in the advanced stages of TB.

However, some literature finds that TB-positive women are more likely to experience a variety of non-specific TB symptoms and less likely to exhibit typical TB symptoms, such as weight loss and night sweats (Breen et al., 2008; Long, Diwan and Vinkvist, 2002). This finding was not supported by a cross-country study in Bangladesh, India and Malawi, where typical male TB symptoms varied across countries (Weiss et al., 2008).

Limitations:

The data are self-reported based on personal experience of symptoms and interaction with the health system. This could imply the presence of reporting errors due to time lapses between interaction with the health system and the survey. Some questions may have elicited incorrect responses due to stigma associated with TB, or TB-specific gender bias. The direction of the possible gender bias is not clear. Traditional concepts of masculinity may prevent men from admitting they are ill and from seeking care (O'Brien et al., 2005). Women, more than men, may under-report TB symptoms due to the possibility of social isolation as a consequence of the disease and associations with HIV/AIDS (Gosoni et al., 2008; Long et al., 2001).

The time lapse between adults' reported interaction with the health system and the timing of the TB cultures may imply that although adults experienced TB symptoms when they first accessed the health system, they did not actually have smear positive TB at that time. Similarly, adults who reported having smear positive test results may have produced negative culture results in the survey due to being put on TB treatment shortly after receiving their smear test results and before the culture test.

The small sample of culture-positive symptomatic adults who sought care at a primary healthcare facility limited my ability to detect statistically significant gender patterns in whether tests were offered to patients and test results.

4.8 Conclusion

While biological sex influences susceptibility to disease, gender has an influence on vulnerability to disease. As gender through vulnerability to disease can be a source of inequity in terms of health outcomes, it is important to understand how it interacts and manifests in specific disease contexts. I explored gender in the context of TB in the Western Cape of South Africa.

In my study, only one third of survey participants indicated they sought help for TB symptoms (coughed ≥ 2 weeks) and only one fourth of those who coughed ≥ 2 weeks reported these symptoms at primary healthcare facilities. It is of concern that some of the TB cases in the sample went undetected. This is not unexpected given that in South Africa TB is diagnosed mainly through passive case finding. However, it is problematic that adults specifically presenting for possible TB symptoms at primary healthcare facilities are not being asked for sputum samples. Almost a third of adults in the sample who presented at primary healthcare facilities with cough symptoms consistent with possible TB were not asked for a sputum sample.

In recent years there has been increasing consideration of active case finding methods to assist in detecting TB (Golub et al., 2005; Getahun and Raviglione, 2010), given the seeming failure of passive case finding methods. This analysis demonstrates that it might be possible to increase the effectiveness of passive case finding through stricter adherence to existing screening protocols at primary healthcare facilities, while also pointing towards the limits of passive case finding. The majority of culture-positive TB cases identified through the survey were individuals who did not report coughing for ≥ 2 weeks or who reported such a cough but did not seek care. This strengthens the call for more active or community-based case finding initiatives.

Women were significantly more likely not to be asked for a sputum sample. This exacerbates the already marginal position of socio-economically vulnerable women who carry a disproportionate HIV burden. I can speculate why women who presented with TB symptoms were less likely than men to be offered sputum tests. Possibly healthcare workers do not view women as the typical TB patient. This possibility needs to be explored and, if so, healthcare workers require sensitisation training to break down existing stereotypes of the typical TB patient. The achievement of gender equity in access to and use of TB services, at the least, requires appropriate implementation of existing health protocols.

Implications for future research:

The mechanism leading to or generating health workers' gender bias in TB testing will have to be identified. Qualitative research in the form of focus group discussions with health workers could aid in understanding whether stereotypes of the gender of the typical TB patient is what is driving the gender difference in TB testing.

Chapter 5

“I didn’t know I was pregnant”: Late access to antenatal care in metropolitan Cape Town

5.1 Introduction

Although maternal mortality cannot be considered an all-encompassing indicator of maternal health⁵¹, it has become a simple and widely accepted signal to understand the performance of countries’ health systems. The WHO (2013) estimated a maternal mortality ratio (MMR) of 300 per 100,000 live births for South Africa for 2010. This made South Africa an extreme outlier with regards to maternal mortality relative to its upper-middle-income country peers. Due to both the direct and indirect impact of HIV and related infections, maternal mortality levels increased rapidly from the late 1990s and only started showing tentative signs of abating since 2010 (Bamford, 2013; Pattinson, Fawcus and Moodley, 2013; Moodley, 2014). However, despite the early signs of decrease in maternal mortality, South Africa’s maternal mortality performance remains concerning given the level of resources available to the health system.

Poor maternal health can have large and enduring impacts on development that are likely to operate through a number of causal pathways. Ill health during pregnancy or birth, irrespective of whether it results in mortality, is associated with lower birth weights for children (Raatikainen, Heiskanen and Heinonen, 2007; Brown et al., 2008) which, in turn, impacts on children’s educational achievement and economic prospects (Currie, 2011; Figlio et al., 2014). Maternal deaths imply the loss of human capital from society, thereby influencing future investment decisions and leading to lower education levels for women (Jayachandran and Lleras-Muney, 2009). Children who grow up without mothers achieve lower levels of education, influencing adult consumption and poverty levels (Beegle, De Weerd and Dercon, 2010; Case and Ardington, 2006; Evans and Miguel, 2007).

A high proportion of births (91.3%) in South Africa occur in healthcare facilities and an even higher proportion (98%) of women have at least one antenatal care visit before giving birth (DHIS, 2012). On average, women in South Africa make 3.7 antenatal care visits during their pregnancy, almost achieving the minimum number of four antenatal care visits recommended by the WHO (DHIS, 2012). However, the majority of women (59.5%) access antenatal care late (≥ 20 weeks/5 months) in their pregnancies (DHIS, 2012). The Department of Health through the District Health Information

⁵¹ It does not account for maternal morbidity which may result in poor birth outcomes for mother and child and also does not include stillbirths.

System (DHIS) measures *early* antenatal care access (<20 weeks/5 months). In this paper I, however, focus on *late* antenatal care access (≥ 20 weeks/ 5 months) as an indicator of the effectiveness of the public health system in reaching women in the critical and vulnerable health state of early pregnancy⁵². The first trimester of pregnancy offers a unique opportunity to ensure that women are sufficiently early in the system to convey the importance of more visits during later stages of pregnancy, and to allow for the detection and eventual treatment of a variety of health conditions which may contribute to preventable morbidity or mortality in pregnancy and birth.

Most importantly, in South Africa it is recommended that, for optimal mother and child health outcomes, HIV-positive women already be initiated on anti-retroviral therapy (ART) by their first antenatal screening or booking visit irrespective of CD4 count (Department of Health, 2014), with evidence that the longer women receive ART before birth the smaller the chances of HIV transmission to the infants (European Collaborative Study, 2007, Hoffman et al., 2010). This is of critical importance in a country where almost a third (29.5%) of women presenting at clinics for their first pregnancy in 2012 was HIV positive (Department of Health, 2013). Almost two thirds (65%) of women whose deaths were attributed to maternal causes in 2011-2013 and 95% of women who died from non-pregnancy related infections (NPRIs), the largest cause of maternal deaths in South Africa, were HIV-positive (Moodley, 2014). Therefore, simply ensuring that the HIV status of all pregnant women is determined early in pregnancy and they are provided appropriate treatment as early as possible, can potentially have a large effect on South Africa's high MMR and child mortality rate.

Given the importance of early antenatal care in minimising the possibility of detectable and treatable negative health outcomes during and due to pregnancy, I measured self-reported late access to antenatal care services through a survey of 221 women in four public health facilities in the metropolitan area of Cape Town (Western Cape), shortly after giving birth. The same survey was used to also explore potential demand- and supply-side explanations for late antenatal care access. I used the same survey instrument as Solarin and Black (2013), who studied the causes of late antenatal care access in the inner-city of Johannesburg in Gauteng province. The study thus builds on the work of Solarin and Black (2013) and allows for the comparison of findings between the metropolitan areas of two large South African cities based in different provinces.

It is unlikely that South Africa's poor maternal health outcomes can be explained purely by the impact of HIV/AIDS or the generally high burden of communicable diseases. The MMR for HIV-

⁵² Using a late rather than early perspective enables a focus on the causes of late access and potentially allows for policy recommendations on how late access can be minimised.

positive women who gave birth in public healthcare facilities is almost six times higher than that of their HIV-negative counterparts (Pattinson, 2012, Black, Brooke and Chersich, 2009). Nevertheless, the MMR for HIV-negative women is still higher than the overall MMR of other middle-income countries with low HIV burdens (Pattinson, 2012). It is, however, important to acknowledge that HIV may affect maternal health outcomes not only as a potential cause of death, but also by directing resources, both in terms of availability and quality, away from general obstetric or antenatal care to other parts of the health system (Graham and Hussein, 2003).

The availability and acceptability of antenatal care services are likely to influence women's perceived value of antenatal care attendance and, ultimately, their willingness to attend antenatal care services. At the same time, antenatal care attendance may also be influenced by knowledge and education levels, access to household and individual resources (which influence direct affordability) and individual experiences of pregnancy.

The field of economics offers useful perspectives to understand why women may delay in accessing antenatal care. Antenatal care is a form of screening or preventative healthcare to ensure the health of the mother and infant later in pregnancy or at birth. Limited access to credit markets or weak savings instruments has been identified as a possible explanation of sub-optimal preventative healthcare consumption (Dupas, 2011). This is an unlikely explanation in South Africa, as antenatal care is offered free of charge in the public health system as part of free access to primary healthcare for pregnant mothers and for children since 1994 and later for everyone (McIntyre et al., 2007). More plausibly, information asymmetries or inconsistent time preferences which lead to individuals placing higher value on health seeking for immediate health conditions (Dupas, 2011) may lead to delayed antenatal care access in South Africa. Furthermore, the importance of sufficient health services in reducing the health seeking burden (Dupas, 2011) should also not be ignored.

The chapter starts with a summary of the determinants of healthcare utilisation, before the definition of maternal mortality and its importance in the global development context is reviewed. This is followed by a discussion of the levels and causes of maternal mortality in South Africa. I then move to considering why minimising maternal mortality and morbidity should be an important policy objective in economic development, beyond pure health considerations, and also review the role of antenatal care in maternal and child health, before discussing the importance of early antenatal care attendance. Once the importance and policy relevance of early antenatal care attendance has been established, an overview of the survey process and structure is provided, before discussing the results of the survey. The chapter concludes by considering the implications of the survey results for South African policy.

5.2 The determinants of healthcare utilisation

Healthcare utilisation and the care seeking which precedes utilisation are influenced by both demand- and supply-side factors. Here a short overview is provided of some of the frameworks which consider factors that influence healthcare utilisation.

In his seminal paper on the demand for health, Grossman (1972) positions health (or good health) as a concept which has both stock and flow components. In his model, people are born with an initial stock of health to which they can add, maintain or allow depletion of during the course of a lifetime. In this simple model, investment decisions in health are, most basically, influenced by age and education levels. Ageing is a process which necessitates greater investments in health over time as health deteriorates with ageing. The shadow price of health, or additional investments required to maintain a stock of health, is asserted to be negatively associated with education, with more educated people having lower shadow prices. The negative association between the shadow price of health and education is based on the assumption that more educated people are more efficient producers of health. In the model, the demand for healthcare is treated as a derived demand and “the shadow price of health depends on many other variables besides the price of medical care” (Grossman, 1972: 225).

While Grossman’s work does not actively consider the healthcare supply-side, there exists a rich literature which explores the factors and frameworks that influence both theoretical and realised healthcare access. This literature tends to focus on the factors influencing access to healthcare, with some frameworks or approaches mainly dealing with potential or theoretical access and others with realised access or actual utilisation (Gulliford et al., 2002) and often draws evidence from developing rather than developed countries. While potential or theoretical access generally takes a supply-side focus, frameworks which are constructed around realised access or utilisation tend to explicitly incorporate the demand-side perspective and take a view of access occurring somewhere in the nexus of the supply- and demand-sides. Non-realised access does not, however, mean that individuals did not have the opportunity to access the system. It is in this regard that McIntyre, Thiede and Birch (2009: 181) view access to healthcare as representing “the *empowerment* of an individual to use health care and reflects an individual’s capacity to benefit from services given the individual’s circumstances and experiences in relation to the health care system”. According to this view, individuals who are empowered may under certain circumstances decide not to access healthcare despite having the opportunity to do so.

Penchansky and Thomas (1981) developed the concept of access into a multi-dimensional measure, consisting of five dimensions, through empirical verification. These dimensions include (Penchansky and Thomas 1981: 128-129):

- *Availability*, or the “volume and type of existing services (and resources) [relative] to clients’ volume and type of needs”;
- *Accessibility*, or “the relationship between the location of supply and the location of clients”;
- *Accommodation*, or “the relationship between the manner in which supply resources are organized to accept clients (including appointment systems, hours of operation, walk-in facilities, telephone services) and the clients’ ability to accommodate to these factors and clients perceptions of their appropriateness”;
- *Affordability*, or “the relationship of prices of services and providers’ insurance or deposit requirements to the clients’ income, ability to pay and existing health insurance”; and
- *Acceptability*, or “the relationship of client’s attitudes about personal and practice characteristics of providers to the actual characteristics of existing providers, as well as to provider attitudes about acceptable personal characteristics of clients”.

In providing definitions for the above access dimensions, both the *nature of the supply side* and the *needs of the demand side* are emphasised (Penchansky and Thomas, 1981). In this framework, it is implied that access occurs at the intersection of these two sides of the market. While the empirical analysis demonstrates that the five dimensions are independent (discriminant validity) and valid (construct validity) constructs, a large degree of overlap remains in the Penchansky and Thomas (1981) framework.

McIntyre, Thiede and Birch (2009) attempted to reduce overlap and simplified the Penchansky and Thomas (1981) framework by collapsing the five access dimensions into three. When reduced in this way, these dimensions include *availability* (also described as physical access), *affordability* (also described as financial access) and *acceptability* (described as cultural access). Acceptability, in this context, is defined as “the fit between provider and patient attitudes towards and expectations of each other” (McIntyre, Thiede and Birch, 2009). These attitudes and expectations can be influenced by factors such as age, gender, ethnicity, race and language. While acceptability has a relatively narrow definition in this framework, availability takes on a wide definition, including “the relationship between the type, range, quantity and *quality* of health care provided at a facility (system factors) and the nature and extent of the health needs of the individuals being served (individual factors)” (emphasis added, McIntyre, Thiede and Birch, 2009: 184).

Goddard and Smith (2001) developed a framework through which to explore equity in access to healthcare in the United Kingdom. They refer to their concept of access to healthcare as an “operationally useful” concept emphasising mainly the supply side, although the demand side is implicitly taken into account by including indirect costs (or opportunity cost) in the framework. This framework distinguishes between the dimensions of *availability*, *quality* (e.g. cleanliness of facilities, time spent with patients), *costs* (direct and indirect) and *information* (do potential users know about the service). In the Goddard and Smith (2001) framework, some aspects (e.g. waiting times and opening hours), which Penchansky and Thomas (1981) describe as accommodation and which McIntyre, Thiede and Birch (2009) classify as availability, are grouped in the quality dimension.

Collapsing many issues related to quality into the availability dimension may lead to a de-emphasis of the importance of quality of healthcare as determinant of health seeking or utilisation, particularly if availability, for ease of operational implementation, is mainly considered in terms of *physical* access. In this chapter, the term acceptability is used to refer to both the quality dimension as identified by Goddard and Smith (2001) and “softer”, cultural perception issues which are likely to influence the behaviour of both clients and healthcare providers, as included under the McIntyre, Thiede and Birch (2009) access dimension of acceptability. The three access dimensions of availability, affordability and acceptability are used in the remainder of the chapter to consider the different factors that influence timely access to antenatal care.

Healthcare access frameworks allow researchers and policymakers to identify the factors which have the biggest influence(s) on healthcare utilisation and ultimately attempt to remedy these factors at a policy level. Not all health seeking or health utilisation behaviour is, however, easily explainable at the level of generalisable factors. After reflecting on a lifetime of work in the area of health economics and health preferences, Feldstein (1995) arrives at the strong conclusion that researchers need to more explicitly account for heterogeneity in health risk preferences and, therefore, also in the healthcare utilisation decisions that people make. He describes this realisation as follows (Feldstein, 1995: 31):

“Although there was no place in my early thinking about the optimal provision of health care for the heterogeneity of tastes and of attitudes towards risk, I now regard that heterogeneity as central. Two individuals with the same income and the same insurance may choose very different medical care just as they choose different lifestyles because of differences in their attitudes toward health risks.”

Lastly, it is important to keep in mind that healthcare utilisation or health seeking behaviour of pregnant women may be influenced by female-specific factors or considerations. This aspect could play out in all of the access dimensions discussed above.

5.3 Maternal mortality and child health in South Africa

Defining maternal mortality:

The WHO defines maternal mortality as “the death of a woman while pregnant or within 42 days of termination of pregnancy, irrespective of the duration and site of the pregnancy, from any cause related to or aggravated by the pregnancy or its management but not from accidental or incidental causes” (WHO, 2015b). The maternal mortality ratio (MMR) is the number of maternal deaths that meet the above WHO definition per 100,000 live births. The MMR took on an especially important role in global development policy, recognising the much heavier burden of maternal mortality on developing countries, when it became one of two targets of the United Nation’s (UN) Millennium Development Goal (Goal 5) relating to maternal health in 2000⁵³. The target set for maternal mortality was to decrease the MMR by 75% between 2000 and 2015 (Blaauw and Penn-Kekana, 2010) and for South Africa this means meeting a target of 38 deaths per 100,000 live births by 2015 (Republic of South Africa, 2013). While it seems that globally the target will be met, South Africa will miss its target (Blaauw and Penn-Kekana, 2010; Bamford, 2012). As discussed later in this section, despite South Africa under-performing relative to its Millennium Development Goal for the MMR, which had been increasing since the late 1990s due to the impact of HIV/AIDS, the MMR reached a turning point in 2009, with clear evidence that it has now started to decrease.

The MMR definition has two aspects which may lead to under-estimation of maternal deaths related to pregnancy or birth: first, the necessity of having to attribute death to the broad (direct or indirect) cause of pregnancy and therefore having to accurately identify a cause of death; and second, timeframe. It is relative to the first aspect that Graham and Hussein (2003) argue that in developed countries, the omission due to misclassification of simply a few maternal deaths may lead to the gross under-estimation of the MMR. In contrast, developing countries not only face the problem of misclassification but also have to contend with weak data systems. Essentially, as emphasised by multiple authors, maternal mortality is a rare event (e.g. Hill, 2001 as cited by Dorrington and Bradshaw, 2011) and this implies that relatively sophisticated data capturing or measurement systems are required. Including direct and indirect causes of obstetric deaths, but excluding

⁵³ The second target of Goal 5 is focused on universal access to reproductive health.

incidental deaths, also creates complications and may lead to missing out on the capturing of deaths that occur at home.

According to the WHO's ICD-10 manual, *direct obstetric deaths* are defined as "deaths resulting from obstetric complications of the pregnancy state (pregnancy, labour and the puerperium), from interventions, omissions, incorrect treatment or from a chain of events resulting from any of the above". In contrast, *indirect obstetric deaths* are defined as "deaths resulting from a previous existing disease or disease that developed during pregnancy and which was not due to obstetric causes, but aggravated by physiological effects of pregnancy" (Dorrington and Bradshaw, 2011: 52).

HIV manifests in direct or indirect causes of obstetric death (Moran and Moodley, 2012; Lathrop, Jamieson and Danel, 2014). In the case of direct obstetric deaths, HIV mainly manifests as pregnancy-related sepsis (Moran and Moodley, 2012). There exists comprehensive evidence that the risk of HIV-infected women developing post-birth sepsis is three times greater than the risk for non-infected women, with this risk increasing to up to six times if women undergo caesarean sections (Calvert and Ronsmans, 2013). For indirect obstetric deaths, HIV substantially increases the chances of death from non-pregnancy related infections (NPRIs) (Moran and Moodley, 2012). In the South African context, the main NPRIs as causes of maternal death are TB, pneumonia and meningitis (Moodley, 2014). TB is the most common cause of maternal mortality amongst all women, HIV-positive and non-infected, in South Africa (Moodley, 2014). In interpreting NPRI statistics, it is important to keep in mind that approximately 90% of women who die from NPRIs in South Africa are HIV-positive (Moodley, 2014).

There also exists growing uncertainty about the appropriateness of a 42-day cut-off for the maternal death definition. While historically the definition may have had its roots in religious practices around confinement and separation of the mother from the family after birth, an alternative (biological) explanation is that women typically only experience their first menstrual bleeding post-pregnancy, 6-8 weeks after giving birth (Høj et al., 2003).

It is increasingly recognised that the impact of pregnancy on a women's health may have consequences that extend beyond the 42-day cut-off. For example, there has been consideration of how suicides that stem from post-natal depression should be treated relative to the definition of maternal mortality (Graham and Hussein, 2003). In recognition of the artificial nature of a 42-day cut-off, the WHO has now adopted a definition of late maternal mortality which extends up to one year after birth and deaths that meet this definition are captured, although not included in the numbers used to calculate the MMR (Graham and Hussein, 2003). While useful to monitor, accurate

data on late maternal mortality would potentially be even more problematic to collect than for the current, narrower definition.

Depending on the availability and quality of data in a particular country, up to three data sources can be used to estimate maternal mortality and the MMR. These include vital registration data (births and deaths), data collected by the health system and population-based surveys (Graham and Hussein, 2003). Population based surveys are particularly useful in contexts where pregnancies, births or deaths that meet the definition of a maternal death occur outside the reach of the health system and also are not captured by other relevant authorities (e.g. the vital registration system). However, these surveys are likely to result in over-reporting if more than one family describe a maternal death and could report deaths outside the 6 weeks cut-off.

Levels and causes of maternal mortality in South Africa:

Given the various data sources that can be used to calculate maternal mortality and South Africa's unique position for a developing country of having of all these data sources (health system data, vital registration data and household surveys) available, the actual maternal mortality statistic for any one year is subject to much uncertainty. South Africa, with input from the relevant data authorities, produces annual estimates. UN agencies (mainly the WHO, United Nations Children's Fund, United Nations Population Fund and the World Bank) have produced joint estimates for global, regional and country-specific maternal mortality every five years since 1990 (Blaauw and Penn-Kekana, 2010). These estimates attempt to control for data quality and the impact of certain diseases such as HIV, which may not be fully accounted for in domestic data. Despite the availability of potentially vastly different domestic estimates, the inter-agency estimates⁵⁴ tend to become the authoritative figures on the South African MMR and are widely cited, including in South African policy documents (Blaauw and Penn-Kekana, 2010). However, since the estimation methodology may be adjusted every five years, the statistics are not comparable over time.

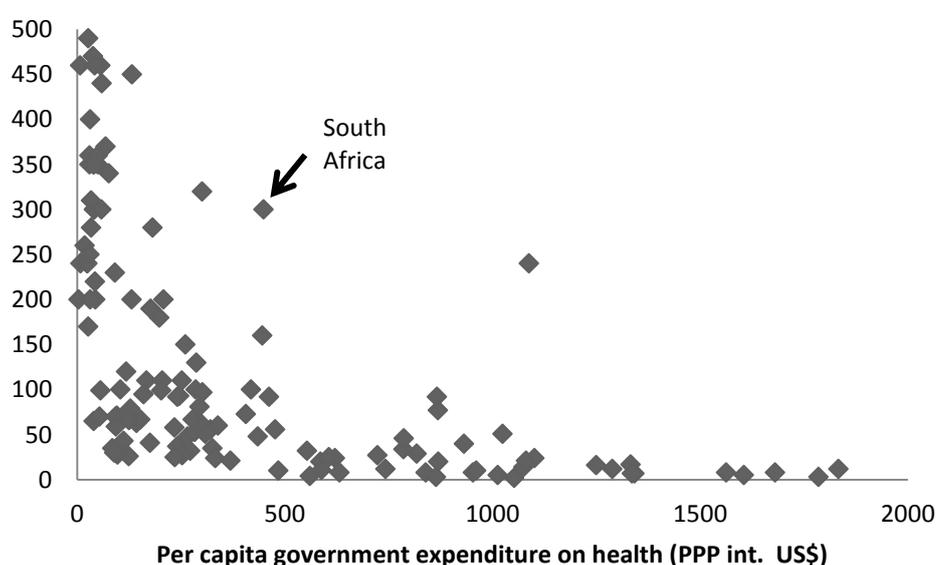
The WHO's (inter-agency) MMR estimate for South Africa for 2010 was 300 maternal deaths per 100,000 live births (WHO, 2013). This made South African an extreme outlier with regards to maternal mortality given its levels of public health expenditure of US\$449.54 in 2011, the closest year to 2010 for which the WHO reports data (WHO, 2013). It also meant that for 2010, South Africa was amongst the forty countries with the highest MMR and shared this position with countries such as Ethiopia, Haiti, Liberia and Niger (see Figure 5.1 below). There is evidence that for many countries the interagency group adds 50% to statistics obtained from the vital registration system to control

⁵⁴ The group of agencies is known as the UN interagency group on maternal mortality estimation (MMEIG) (Dorrington and Bradshaw, 2012).

for under-registration of maternal deaths (Dorrington and Bradshaw, 2012). This may have artificially inflated South Africa's MMR for 2010. New inter-agency estimates are generated and previous estimates are updated every 3-5 years (WHO, 2015b). The 2010 MMR estimate has been updated following the most recent round of inter-agency MMR estimates and the MMR provided for South Africa for 2010 is now 140 maternal deaths per 100,000 live births (WHO, 2015c). The estimate for the most recent year available (2013) is also 140 (WHO, 2015c).

Figure 5.1: Maternal mortality ratio of various countries (2010) relative to per capita government expenditure on health (PPP int. in US\$ for 2011)

Source: Author's own calculations using data from WHO (2013)



In 2010 Blaauw and Penn-Kekana (2010: 10) concluded that “at present, the best source of data on the causes of maternal deaths in South Africa is the assessment conducted by the National Committee for Confidential Enquiries into Maternal Deaths (NCCEMD)”. However, in 2011 the Health Data Advisory and Co-ordination Committee (HDACC), which is responsible for the quality and integrity of health data in South Africa, recommended the best source for maternal mortality data is vital registration data which has been adjusted to control for under-registration and misclassification of causes of death (Dorrington and Bradshaw, 2012; Bamford, 2013).

The NCCEMD produces the “Saving Mothers” report on maternal mortality in South Africa, in which it reports on MMR levels and causes of maternal deaths as identified through the Confidential Enquiry process. The Confidential Enquiry process started in October 1997, with the first report published in October 1999 for deaths which occurred during 1998 (Pattinson, 2012). Data from the District Health Information System (DHIS) is used to calculate the denominator (total births), while

direct submissions to the NCCEMD via provincial governments are used to obtain data on maternal deaths (Pattinson, 2012)⁵⁵. Reports produced from the data provide a summary of *institutional* maternal deaths and their causes for periods of three years (trienniums), but interim reports providing data on maternal mortality for shorter periods are released before the larger (triennium) reports are available.

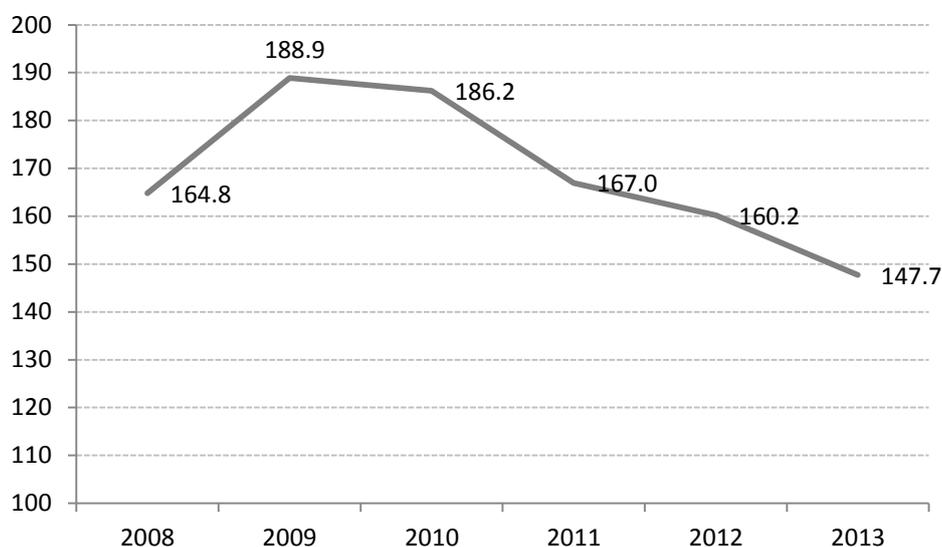
In interpreting the findings of the NCCEMD, it is important to keep in mind that it only provides information on maternal death levels and causes of deaths within the public healthcare system and where cause of death could be linked back to the health states of pregnancy and/or birth. While the data it collects on the number of deaths may be useful in triangulating maternal mortality figures captured by the vital registration system, its value to policymakers lies primarily in the confidential nature of the enquiry and the honesty it therefore elicits from health workers in terms of reported causes of maternal death. All data on reported deaths are also destroyed after the completion of each triennium report (Pattinson, 2012). The extremely confidential approach of the Enquiry is what allows it to “describe the magnitude of the problem of maternal deaths, the pattern of disease causing maternal deaths, the avoidable factors, missed opportunities and substandard care related to these deaths and...[make] recommendations concerning ways of decreasing maternal deaths” (Pattinson, 2012: 1).

The most recent Saving Mothers Report (2011-2013) reports a turning point in maternal mortality. While tentative signs of maternal mortality growth abating were observed in 2010 data (Pattinson, Fawcus and Moodley, 2013), the 2011-2013 report and the institutional maternal mortality ratios (iMMRs) for these years confirmed a definitively decreasing trend (Moodley, 2014) (Figure 5.2). The large-scale roll-out of ART to pregnant mothers which has taken place since 2009/10 is likely to be a large contributing factor to decreasing maternal mortality ratios. Prior to 2010, only pregnant women with AIDS defining criteria qualified for ART (Bamford, 2013). This changed to ART initiation for HIV-positive women with a CD4-count of 350 cells/ μ L or less during 2010 (Department of Health, 2010). Since April 2014, all pregnant and breast feeding women qualify for ART regardless of their CD4 cell count. (Department of Health, 2014). It is likely that further improvements in the MMR can be expected from this latter change.

⁵⁵ A comparison between data obtained through the NCCEMD process and maternal deaths recorded in the DHIS shows that the DHIS is under-capturing maternal deaths by almost a third, while some health districts that report into the DHIS are not reporting maternal deaths to the NCCEMD (Pattinson, Fawcus and Moodley, 2013).

Figure 5.2: Recent institutional maternal mortality ratios (iMMRs)⁵⁶

Source: Individual year figures (unrevised) as identified by the Confidential Enquiry into Maternal Deaths and accessed on Health Indicators page of Health Systems Trust (24 April 2015)



The most recent comprehensive *triennium* Saving Mothers Report (2011-2013) found that non pregnancy-related infections (e.g. HIV and TB) is the largest cause (34.5%) of maternal deaths⁵⁷ - 90% of women who died from this cause were identified as HIV-positive (Moodley, 2014). Obstetric haemorrhage (15.8%) and hypertension (14.8%) were respectively the second and third biggest causes of maternal death (Moodley, 2014). Collectively the top three causes are responsible for two thirds of maternal deaths in South Africa. While the iMMR of non-pregnancy related infections and hypertension as causes of maternal death decreased from the previous triennium Saving Mothers Report (2008-2010), the iMMR of obstetric haemorrhage as cause of death has increased (Pattinson, 2012; Moodley, 2014). Together the top three causes contributed 65% of avoidable maternal deaths, with maternal deaths attributed to obstetric haemorrhage and hypertension-related conditions identified as being “possibly and probably preventable” for the majority of these cases.

The recent report notes a large increase in HIV-positive women who died from non-pregnancy related infections and who were receiving Highly Active Antiretroviral Therapy (HAART): 55% were receiving HAART compared to only 36% in 2008-2010 (Moodley, 2014). There are various possible reasons why these women died from non-pregnancy related infections despite receiving HAART (Moodley, 2014). Nevertheless, the important point is that the increase in pregnant women

⁵⁶ Apart from a single data point for 2006, the iMMR was not reported on for individual years before 2008. The NCCEMD triennium reports before the 2008-2010 report only provided an average iMMR for each three-year reporting period.

⁵⁷ The report describes non-pregnancy related infections as “mainly deaths due to HIV infection complicated by tuberculosis (TB), PCP and pneumonia” (Moodley, 2014).

receiving HAART amongst those who died from non-pregnancy related infections most likely implies a similar or larger increase of pregnant women who received HAART and remained healthy.

Significantly, the report draws strong conclusions about the role and quality of maternal healthcare provided, especially for avoidable maternal deaths. It is argued that for the 60% of total maternal deaths classified as “possibly and probably preventable”, “poor quality of care during the antenatal, intrapartum and postnatal periods” is likely to have played a large role (Moodley, 2014: vi).

5.4 The benefits of healthcare that minimises maternal morbidity and mortality

The potential impacts of maternal mortality and morbidity (or sub-optimal health) on economic development operate through a variety of channels. The death of a mother not only removes a potentially productive individual from society implying a loss of human capital, but also robs her infant (if the infant survives) and any other remaining children of a mother. Furthermore, adverse health conditions during pregnancy can have negative impacts on the health state of the infant at birth, meaning that such children start life with a health deficit relative to other children. Health deficits at birth can have enduring effects on socio-economic status and mobility prospects during the lifetime of children. Here I provide a brief review of available evidence on some of the channels through which the impact of maternal mortality or ill health may operate.

There is both individual country and cross-country evidence available on the relationship between maternal mortality and human capital accumulation. Jayachandran and Lleras-Muney (2009) consider the effect of a sharp decrease in maternal mortality in Sri Lanka between 1946 and 1953 on education investment, which leads to human capital accumulation over the longer term. They argue that lower maternal mortality is likely to influence education investment through an increase in life expectancy for women and a rational updating of investment choices, given longer lives for women. They exploit district-level variation in the MMR and find that by age 15 the decrease in the MMR had led to a 2.5% increase in female literacy and a 4% increase in years of schooling for women.

Albanesi (2011) considers the relationship between maternal mortality, fertility levels and human capital accumulation in 25 developed and emerging market countries in the period 1900-2000. She finds that while maternal mortality declines initially led to increased fertility (periods of so-called “infant booms”), eventually fertility levels started decreasing and many of these economies experienced “infant busts” or periods of fertility decline. More specifically, she finds that “infant busts were more pronounced in countries with high income per capita, high proportion of industry and high levels of secondary education”, being suggestive of a positive relationship between infant busts and human capital accumulation (Albanesi, 2011: 3). This relationship is confirmed through

empirical results that reveal an increase in the female-male education ratio in countries which experienced large maternal mortality decreases.

Increasingly it is recognised that the health state of an infant at birth is no longer simply the result of unchangeable genetic factors, but more likely the result of a variety of socio-economic and environmental conditions and behaviours interacting with genetic factors (Currie, 2011). Moreover, the economics literature has established that health state at birth is a strong determinant of future socio-economic outcomes such as earnings and educational achievement (Currie, 2011).

Access to care, or the absence thereof, during pregnancy could influence the health state of both the mother and the child. Under- or non-attendance of antenatal care is associated with lower birth weights for infants (e.g. Raatikainen, Heiskanen and Heinonen, 2007; Brown et al., 2008). Figlio et al. (2014) study the impact of neonatal health, as measured by birth weight, on cognitive development and outcomes for a large sample of children born in Florida between 1992 and 2002. They find neonatal health affects cognitive development during childhood and that this relationship is present in a number of different sub-groups of children (i.e. twin comparisons, sibling comparisons and lone children or “singletons”). Furthermore, when the authors control for school quality in regressions to establish the relationship between birth weight and cognitive development, they find no effect on regression results from the inclusion of school quality. They therefore conclude that the school system is an unlikely enabler for catch up in cognitive development for those children born with a neonatal health deficit.

The last channel through which maternal mortality could impact economic development is effects on other remaining children, or even on the life of the infant that has to grow up without a mother. It is estimated that globally there are 17.8 million AIDS orphans, growing up without a mother, father or even both parents, of whom 85% reside in sub-Saharan Africa (UNICEF, 2013). At least three relatively recent studies (Beegle, De Weerd and Dercon, 2010; Case and Ardington, 2006; Evans and Miguel, 2007) use panel or other longitudinal data to study the effect of orphanhood in specific African countries on several measures related to schooling. All three studies find a much stronger effect of maternal death on children’s educational achievements compared to the effects of paternal death.

Case and Ardington (2006), using longitudinal data for a sample of South African Zulu-speaking children, find that there is a strong and significant association between maternal death and not being enrolled at school or completing fewer years of schooling compared to their non-orphaned counterparts with whom they live. Furthermore, conditional on being enrolled, maternal orphans

receive lower level of expenditure on their schooling. In contrast, no such effects are found for paternal orphans, although there is a large and significant association between lower household socio-economic status and having lost a father. The authors argue that the absence of a father, a common situation for many South African children irrespective of their orphanhood status, may help to explain the strong maternal orphan effect as “father absence turns maternal orphans into virtual double orphans” (2006: 419).

Evans and Miguel (2007) use panel data for a sample of Kenyan children. They measure impacts on school participation, or the percentage of days on which children were at school when unannounced enumerator visits took place. There is a large decrease in school participation following death of a parent. The largest impacts are found for children who lost a mother rather than a father, while larger impacts are also identified for children who were not performing well at school before the loss of a parent.

Lastly, Beegle, De Weerd and Dercon (2010) use panel data to study the impacts of orphanhood on schooling attendance and height outcomes of Tanzanian children. The loss of a mother reduces total educational attainment by one year and also leads to children achieving a final height of 2cm less than non-orphaned children. These health and education losses imply an overall reduction in per capita consumption of 8.5%. No causal effects are found for paternal orphans.

5.5 Antenatal care and maternal and infant health outcomes

The benefits of routine antenatal care attendance for both mother and infant health has been established in medical literature (see e.g. Carroli et al., 2001; Kulier et al., 1998). Antenatal healthcare visits not only matter in terms of more direct benefits like health services provided during visits, but also more indirect benefits such as providing women with information on pregnancy dangers signs which may enable faster responses to pregnancy health emergencies (Department of Health, 2007) or making women more likely to deliver in a healthcare facility (Bloom, Lippeveld and Wypij, 1999). Medical literature emphasises the health benefits of providing health services like screening and treatment for diabetes and gestational diabetes, screening for preeclampsia through blood pressure measurement and treatment through calcium supplements and screening for iron deficiency or anaemia which can then be treated through iron supplements (e.g. Banta, 2003). Except for a few studies (e.g. Bloom, Lippeveld and Wypij, 1999), the indirect benefits of antenatal care have not received much attention.

It is, however, important to caution that while the literature informing the above discussion and conclusions draws on the results of randomised control studies, much of the medical literature on

the benefits of antenatal care relies on descriptive studies and statistical associations rather than causal analyses (Conway and Kutinova, 2006). In the economics literature, researchers have mainly focused on benefits to the infant rather than the mother and have in terms of infant benefits mainly focused on birth weight due to ease of measurement (Conway and Kutinova, 2006). Despite various searches using different keywords, only two economics studies were found that employ causal estimation methods to consider the benefits of antenatal care for maternal health, namely Conway and Deb (2005) and Conway and Kutinova (2006).

Conway and Kutinova's (2006) main contribution is to widen the economics literature by demonstrating that antenatal care provides health benefits that extend to the mother and not only the infant, particularly if maternal health benefits are defined in terms of less studied and more subtle measures. Using the National Maternal and Child Health Survey from the US, Conway and Kutinova use instrumental variables for maternal postnatal weight and the post-birth length of hospital stay. They find that accessing "timely and adequate" antenatal care increases the possibility of a healthy postnatal weight for the mother and, for the more socio-economically vulnerable group of African-Americans in their sample, that antenatal care decreases the probability of long hospital stays after birth.

Conway and Deb (2005) argue that the benefits of antenatal care may be under-estimated if data on complicated and uncomplicated pregnancies are combined. They argue that a relatively small percentage of pregnancies (10-15%) experience complications and that, for these pregnancies, antenatal care is typically delivered through a large number of visits which may have little impact on birth outcomes due to the nature of the pregnancy. In contrast, antenatal care may have a far larger impact on uncomplicated pregnancies and it is therefore important to study these two groups of pregnancies separately. Using data from the US National Maternal and Child Health Survey, they separate complicated and uncomplicated pregnancies and apply a finite mixture model to these two groups of pregnancies to consider the impact of onset of antenatal care (when antenatal care was first accessed) on infant birth weight. While they find no impact for complicated pregnancies, timely antenatal care for uncomplicated pregnancies increases infant birth weight by 30-35 grams for each week antenatal care was accessed earlier.

Apart from general attendance and access to healthcare during pregnancy, antenatal care has three dimensions that potentially matter for maternal and infant health outcomes: 1) the frequency of antenatal care visits; 2) timing of access to antenatal care; and 3) the type of services or the quality of antenatal care provided. Variation in these dimensions may influence the perceived or actual

benefits of antenatal care. Below I provide some provide brief evidence from the medical literature of how these three dimensions matter for health outcomes and health system engagement.

Timing:

The issue of timely or early antenatal care attendance has been relatively under-emphasised in both the health and economics literatures, with it only recently receiving more attention due to the prominence of HIV as a disease that causes maternal health complications. One of the only medical studies identified that focuses directly on the association between health benefits and early antenatal care attendance is from Finland, most probably due to the availability of sophisticated health system data in that country. In this study (Gissler and Hemminiki, 1994), accessing antenatal care services after 16 weeks of gestation was significantly associated with lower birth weight, giving birth before 37 weeks of gestation, infants requiring care in neonatal units shortly after birth and obtaining a lower APGAR score (a scoring system used to rate the infant's health shortly after birth).

The results of the study by Conway and Deb (2005), discussed above, also provide evidence of the importance of timely access to antenatal care. For uncomplicated pregnancies, they found that infant birth weight increases by 30-35 grams for each week that antenatal care was accessed earlier.

Indirect evidence on the role of timing of antenatal care visits in health outcomes is provided by South Africa's Confidential Enquiries into, respectively, Maternal and Perinatal Deaths. The Enquiry identified late booking (antenatal care attendance) as a top three probable patient-avoidable factor for all major causes of perinatal deaths (infants weighing 1,000 grams or more) and as accounting for 16.8% (1432/8502) of probable patient-avoidable factors in all causes of death for infants weighing 1,000 grams or more. Based on assessment of causes of death and identification of patient-avoidable factors in these deaths, the 2008-2010 NCCEMD recommended booking before 20 weeks of pregnancy for identification of HIV and high-risk medical conditions. (Pattinson, 2012).

Most importantly, in the South African context, early access to antenatal care matters for HIV screening and the timing of the initiation of ART if women test HIV-positive. South Africa's current PMTCT guidelines recommend that HIV-positive women be initiated on ART therapy at their first antenatal care or full screening (booking) visit, irrespective of their gestational age or CD4 cell count (Department of Health, 2014). In the period 2010-2013 South Africa implemented the WHO's "Option A" according to which all HIV-positive mothers were to receive zidovudine (AZT) as a form of ARV prophylaxis from 14 weeks of gestation, following which their CD4 count would be tested and full ART initiated depending on CD4 count level (Schnippel et al., 2015). The mother's total period of ART exposure influences the risk of transmission of HIV to the infant *in utero*, during birth and the

breastfeeding period (European Collaborative Study, 2007, Hoffman et al., 2010). The fact that non-pregnancy related infections is the major maternal cause of death and that the majority of these women (95%) are HIV-positive (Moodley, 2014) can be considered indicative of immune suppression which may have been minimised if ART had been initiated sufficiently early in pregnancy. Given that delays in initiating patients on ART due to health system rather than patient factors have been documented in South Africa (Myer et al., 2012; Schnippel et al., 2015) and have been found to not improve health outcomes for HIV-positive mothers (Myer et al., 2012), it is prudent that women access the system as early as possible, ideally as early as the first trimester of pregnancy.

Frequency:

In Finland, under-attendance (only 1-5 antenatal care visits compared to sample average of 17) is associated with low birth weights, more fetal deaths (death before 22 weeks of gestation) and more neonatal deaths (Gissler and Hemminki, 1994). In Kenya, women attending antenatal care at least twice before giving birth were less likely to have stillbirths (Brown et al., 2008).

Earlier access to health care is likely to allow for a longer time period for more antenatal care visits before birth. However, it is important to keep in mind that beyond a certain threshold more visits may be associated with higher risk pregnancies requiring more monitoring.

Quality:

Despite early access to antenatal care being important for a number of potential health outcomes, it is not only access and frequency that matter but also the content and nature of visits. Improved health outcomes can only be realised if women are provided with the health services that matter for these outcomes. The analysis of health system records for a sample of pregnant mothers attending antenatal care in rural Kwazulu-Natal revealed that these women were infrequently provided with all components of essential antenatal care services (Hoque, Hoque and Kader, 2008). Only 18% of women were tested for syphilis at their first visit and only 16% of women who tested positive for syphilis received treatment. Furthermore, at 36 weeks of gestation only 2% of women were given a full term pregnancy syphilis test, while only 55% of women were given the full three doses of tetanus vaccination before giving birth. A study on the timing of the initiation of antenatal care attendance in inner-city Johannesburg also found low levels of compliance on some service components considered antenatal care best practice (Solarin and Black, 2013). In contrast to the Kwazulu-Natal study which used health system records, the latter study relied on women's self-reported experiences of services provided at the first booking or screening visit which may generate under-reporting due to recall bias. Only 67.2% of women reported having been informed of

pregnancy danger signs, while only 59.6% reported having all routine health checks included in the survey done at the first screening visit.

Women's experiences of the content and nature of antenatal services are likely to influence their willingness to remain in health care and return for further care. In Uttar Pradesh, India women were more likely to use safe delivery options (birth in a healthcare facility or attended by a health professional) if they received better quality antenatal care services, as measured by an index that included measures for timing, frequency and actual services provided (e.g. whether blood pressure was measured) (Bloom, Lippeveld and Wypij, 1999).

5.6 Factors associated with late antenatal care attendance

The topic of late antenatal care attendance has been explored by several South African and international studies that use qualitative and quantitative approaches to identify the correlates of late (or early) antenatal care attendance. Detailed quantitative and specifically causal analyses on factors influencing antenatal care attendance are rare. Most of the studies tend to have small sample sizes – one of the qualitative studies referred to in this discussion had a sample of only 12 women (Ngomane and Mulaudzi, 2010). The definition of late attendance varies across studies – some define it as attendance after 12 weeks gestation or the first trimester (e.g. Finlayson and Downe, 2013), while others use a definition of attendance at 20 or more weeks gestation (e.g. Haddril et al., 2014; Solarin and Black, 2013). Here a brief review is provided of the main categories of correlates that have been identified from both quantitative and qualitative work by placing these factors in an access framework. In this framework factors related to affordability, availability and acceptability are considered. The section is concluded by providing a high-level summary of the main findings of the Solarin and Black (2013) study that provides the methodological approach for this study.

Healthcare availability:

Healthcare availability refers to the physical and geographic availability of healthcare services. This includes the type and quantity of services offered (McIntyre, Thiede and Birch, 2009).

Healthcare availability in the physical access sense is infrequently emphasised as a key factor constraining early access, with most studies focusing on other, often softer factors. In a review of literature on factors constraining timely access to antenatal care, Finlayson and Downe (2013) identify “difficult and dangerous” travel as a sub-theme related to both limited personal and health system resources. The clearest evidence of supply-side constraints in the delivery of care is provided by Solarin and Black's (2013) study in inner-city Johannesburg where 49% of women in their study

reported not receiving a full screening or booking visit the first time they attended at the antenatal care clinic.

Healthcare affordability (direct and indirect costs):

Affordability refers to the direct and indirect costs of healthcare relative to the incomes of users. Antenatal care in the public system is offered free of charge but indirect costs may apply. These indirect costs are likely to have a larger effect on health seeking behaviour if users have lower incomes.

Socio-economic status appears as an important correlate of delayed antenatal care access. A number of qualitative studies assert that women with seemingly lower socio-economic status or who can be considered relatively deprived are more likely to access antenatal care later in their pregnancies (Downe et al, 2009; Haddril et al., 2014). Data from the South African National HIV Prevalence, Incidence, Behaviour and Communication Survey (2008) show a large disparity in antenatal care attendance rates before 20 weeks of gestation for women in the poorest quartile compared to women in the richest quartile (Wabiri et al., 2013). There also exists quantitative evidence from a South African study that rural women with lower socio-economic status, as measured by an asset and access to services index, use hospital maternal delivery services at a lower frequency than one would expect given the prevalence of pregnancies amongst this group (Silal et al., 2014). In the South African context, where education levels tend to be predictive of socio-economic status, a significant positive association between late antenatal care attendance and not completing secondary education has also been found (Solarin and Black, 2013). While these studies identify socio-economic status and lower education levels as correlates of delayed attendance, they do not employ causal analysis techniques. The studies most likely are subject to simultaneity, with poverty and lower education levels being correlated with other factors that also influence delayed attendance.

In the context of a constrained antenatal healthcare supply-side, Solarin and Black (2013) find a positive and significant association between being employed and late attendance (Solarin and Black, 2013). This can be considered a reflection of the opportunity cost of antenatal care attendance if attendees have to miss work and, in some case, forego income as reported in other studies (Abrahams, Jewkes and Mvo, 2001).

Healthcare acceptability:

Acceptability refers to the various factors that influence users' and healthcare workers' perceptions of each other and perceptions of the quality of healthcare services.

Women's negative experiences of antenatal care and their experiences of the actual (or perceived) value that it offers are identified as a common theme by a number of studies (Abrahams, Jewkes and Mvo, 2011; Downe et al., 2009; Finlayson and Downe, 2013; Myer and Harrison, 2003). In resource constrained settings, women may experience long waiting times and poor access to equipment and technology in the delivery of the service (Finlayson and Downe, 2013). Health staff rudeness and abuse could also contribute to a negative experience of antenatal care (Finlayson and Downe, 2013). Limited communication and information flows between staff and women who attend antenatal care may also contribute to the experience that antenatal care offers limited value. After semi-structured interviews with a sample of late attending women in the Cape Town public health sector, Abrahams, Jewkes and Mvo (2001: 246) conclude:

"...the women's own needs were never met by the services they received. They were given very little information regarding when they will deliver, the status of the infant after each palpation and test results; this, undoubtedly, contributed to their ambiguous views on the value of antenatal care. Attendance might, therefore, be improved if women's own perceived needs for information were met."

Other factors:

Late attendance may also be associated with engagement in other risky behaviour such as substance abuse (drug or alcohol use) and smoking. Downe et al. (2009) identify this issue under a theme referred to as "influence of chaotic lifestyles" and highlight the self-reported lack of planning in the life of an alcohol-dependent woman.

Fear of the results of an HIV test in the context of high HIV prevalence rates in South Africa and where an HIV test is offered to women as part of a series of blood tests may also play a role in delayed access. Stinson and Myer (2012) mention that women often do not test for HIV before pregnancy and have difficulty accepting news of an HIV-positive status once diagnosed at or following their booking visit. Amnesty International (2014) identified fear of HIV testing as a factor affecting antenatal care access in South Africa.

Haddrill et al. (2014) identifies the theme of "not knowing" after semi-structured interviews with a sample of late antenatal care attenders in the United Kingdom. These women did not identify the

symptoms of pregnancy or were unable to link physical signs or experiences to the realisation of pregnancy (Haddril et al, 2014). This theme is common to most studies which identify factors associated with late attendance, also in South Africa (Abrahams, Jewkes and Mvo, 2001; Solarin and Black, 2013).

The number of pregnancies a woman has experienced is also likely to influence the timing of antenatal care attendance. It may be more difficult to recognise pregnancy symptoms with a first pregnancy (nulliparous women), leading to delayed attendance (Downe et al., 2009), or woman who have experienced pregnancy before (multiparous women) may postpone attendance if earlier pregnancies were problem-free (Finlayson and Downe, 2013). On the other hand, complications experienced during previous pregnancies could be associated with earlier antenatal care attendance (Abrahams, Jewkes and Mvo, 2001).

Denial of the pregnancy once it is discovered, especially in situations where life or socio-economic circumstances make pregnancy less than ideal, is mentioned by several studies (Downe et al., 2009; Haddrill et al., 2014). An unplanned pregnancy may lead to denial or negative emotions at the time of pregnancy discovery and influence any subsequent health actions (Downe et al., 2009; Stinson and Myer, 2012).

Women could also avoid seeking antenatal care before the pregnancy is physically confirmed (before they are “showing”) (Stinson and Myer, 2012). In the South African and other developing country contexts, women report not wanting to engage in behaviour that confirms pregnancy because of a fear of bewitchment (witchcraft) (Ngomane and Mulaudzi, 2010) or possible “cursing” (Finlayson and Downe, 2013) of the pregnancy or unborn infant. In addition, some women mention experiencing pregnancy as a positive health state which does not necessitate healthcare (Myer and Harrison, 2003; Finlayson and Downe, 2013).

Findings from Solarin and Black (2013):

I used the study design and survey instrument from the Solarin and Black (2013) study as basis for the survey. Some amendments to the instrument were made and these changes are described in Section 5.7. Given the overlap in study design and methodology, a relatively detailed summary of the Solarin and Black (2013) study findings is provided before proceeding to a description of the survey.

Solarin and Black (2013) explored the issue of late antenatal care access in the context of three public healthcare facilities in inner-city Johannesburg. They did this by conducting interviews with 208 women at three public sector labour wards shortly after giving birth during October 2009. The sample size is representative of the size of the facilities given the total number of women who gave

birth there during the year preceding the survey. Based on self-reported timing of first antenatal care attendance, they find that while 97% of women sought antenatal care at least once before birth, only 46% sought care before 20 weeks of pregnancy.

While a large proportion of women sought healthcare late because they were unaware that they were pregnant (21.7%), were subject to time constraints (20.8%) or elected to rather seek care from a private sector general practitioner (17.0%), supply-side factors also played a large role in antenatal care delay. In antenatal care terms, the “booking” visit is defined as the visit at which the first antenatal screening is conducted, e.g. a patient file is opened, a test is conducted to confirm her pregnancy and a series of health checks, including an HIV test and blood pressure measurement, is conducted. Of the 198 who sought care in the sample, 97 did not have a booking appointment at their first visit, i.e. they were not taken in by the clinic for a full screening visit. Of this group 39.2% experienced a delayed booking visit after their first clinic visit because clinic staff told them to return more than a month later. The median gestational (pregnancy) age of the latter group of women at their first visit was 4 months, while the median gestational age at their booking visit was 7 months, implying a 3 month delay in having a “booking” visit at the antenatal clinic after being told to return later.

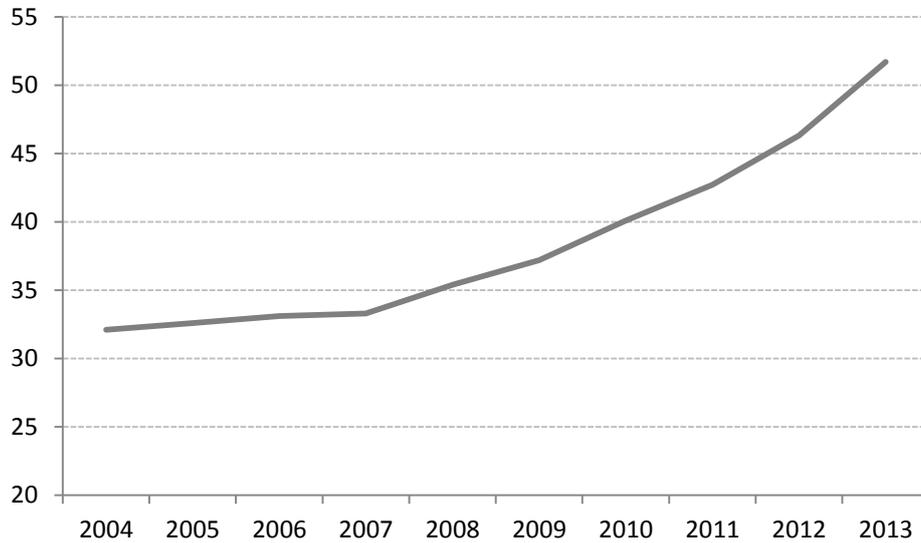
The study identifies multiple reasons for a delay in the booking visit after the first clinic visit. The study also finds that 7.1% of women who were asked to return on another day had to do so between 3 and 6 times before their booking visit took place.

Magnitude of late antenatal care attendance in South Africa:

The South African Department of Health monitors the timing of antenatal care attendance through the District Health Information System used for the reporting of administrative data by all levels of the public health system. The percentage of mothers that have their first antenatal care visit before 20 weeks of pregnancy is calculated relative to the total number of first antenatal care visits (denominator) (District Health Information System definition as reported in Day and Gray, 2014). The government has identified an increase in early access to antenatal care services to more than 60% as an explicit goal to be achieved by 2016 (Department of Health, 2012). Early attendance rates have increased at an accelerated rate since 2008, with the national average early attendance rate totalling 51.7% in 2013 (Figure 5.3).

Figure 5.3: South African national rate of antenatal care attendance at public health clinics before 20 weeks gestation

Source: District Health Information Service (DHIS)

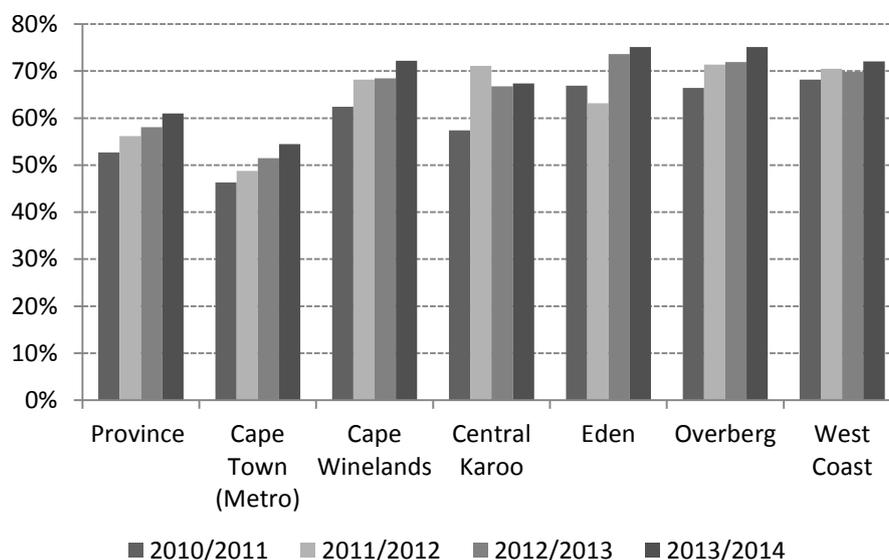


5.7 Study context

The number of women accessing antenatal care before 20 weeks of pregnancy in the Western Cape has been slowly increasing over the last few years (Figure 5.4). Despite these increases, the low percentage of women accessing antenatal care before 20 weeks of pregnancy remains an area of concern to Western Cape policy-makers and is highlighted as such in the province's Department of Health annual performance plans (Western Cape Government: Health, 2013-2015a).

Figure 5.4: Western Cape rate of antenatal care attendance before 20 weeks gestation

Source: Data from various Western Cape Government: Health, Annual Performance Plans



Metropolitan Cape Town is the worst performing health district in the Western Cape in terms of antenatal care attendance before 20 weeks gestation (Figure 5.4). Given that this district was responsible for between 64.7% and 66.0% of total first antenatal care visits⁵⁸ over the reporting period 2010/2011 to 2013/2014, low levels of antenatal care access in this district have a large impact on the province's overall performance in the area of early access to antenatal care.

Several policy-makers were met as part of a consultative process on the policy relevance of the research and also to understand whether specific districts or sub-districts should be targeted in the research. Some policy-makers attributed part of the Metropolitan district's weak performance in early access to antenatal care to historical difficulties in coordinating the delivery of antenatal care between the Western Cape Department of Health (WCDOH, provincial government) and City of Cape Town (local government) health services. WCDOH integrates antenatal care with other primary healthcare services at primary healthcare facilities (clinics and community healthcare centres), while the City of Cape Town offers antenatal care services at some of its own primary healthcare facilities. Part of the historical coordination failure between the two levels of government was that women were able to only test for pregnancy at City of Cape Town clinics and, if identified as pregnant, were then referred to WCDOH clinics for antenatal care. Through this referral process, loss to follow up occurred and some women did not return for antenatal care or only returned much later in their pregnancies. Antenatal care services, however, are now being offered at City of Cape Town clinics.

⁵⁸ The denominator used to calculate antenatal care access before 20 weeks.

Given the relatively poor performance of the Metropolitan district, two sub-districts were selected for the survey. Within each of these two sub-districts, the two main labour wards to which women are referred for birth by their local clinic offering antenatal care services were selected— one district hospital and one maternity obstetric unit (MOU) based in a community health centre (CHC). The sub-districts were selected due to several reasons. These sub-districts were relatively under-studied from a health perspective compared to other sub-districts in Metropolitan Cape Town and also were based in relatively safe and accessible environments for ease of access of interviewers. Sub-districts that had specific interventions to increase the early antenatal care attendance rate, such as community healthcare worker programmes aimed at pregnant women or women with young children, were avoided.

According to information on the Western Cape Department of Health's website (2015), antenatal care services for low-risk pregnancies are provided at midwife obstetric units in urban areas and fixed or satellite clinics in rural areas. Women with high-risk pregnancies receive services from outpatient clinics at urban and rural hospitals once their pregnancy has been categorised in terms of risk. Women are encouraged to book before 20 weeks and it is indicated that they need to attend an antenatal care visit every 6 weeks following their booking visit, up to 28 weeks gestation. Thereafter the next visit will be at 34 weeks and then visits will be scheduled according to their individual pregnancy profiles. At the booking visit, they are required to bring their identity document, their clinic or hospital card if they had previously attended the same facility and also information on any medication they are currently taking (Western Cape Government: Health, 2015d).

The survey interviews were conducted between 20 October and 20 November 2014 after pilot testing the survey at one facility. All collected data were captured in Epidata 3.1 and analysed using Stata 13.

Health system indicators:

Health system indicators captured by the District Health Information System provide a snapshot summary of both the disease burden and the South African health system's ability to effectively manage and improve the disease burden at district level. While these indicators cannot be considered pure performance measures as they are influenced by a number of factors outside the control of the health system, they are indicative of the relative performance of health districts or sub-districts. Table 5.1 (below) presents a number of maternal and child health indicators for both the district and sub-district in which the Solarin and Black (2013) study was conducted and the two health sub-districts of this study.

Table 5.1: Maternal and child health indicators for study sub-districts⁵⁹

Source: DHIS, National Department of Health

	Johannesburg Metro ⁶⁰ (2009)	Johannesburg Region F (2014)	Cape Town Tygerberg sub-district ⁶¹ (2014)	Cape Town Eastern sub-district (2014)
Antenatal 1st visit before 20 weeks rate (rate)	21.5	44.8	64.2	60.8
Antenatal client HIV 1st test positive (rate)	26.5	15.2	4.0	5.4
Antenatal client initiated on ART (rate)	43.4	92.5	60.3	77.8
Maternal mortality in facility ratio (annualised)	Not available (incomplete data capturing)	116.7	63.1	19.0
Neonatal mortality in facility rate (annualised)	10.5	32.1	14.6	0.5
Child under 5 years pneumonia case fatality (rate)	Not available	6.1	1.6	0.7
Child under 5 years diarrhoea case fatality (rate)	8.8	8.9	0.3	0.4

The indicators for Johannesburg Region F for 2014 have to be interpreted relative to the large-scale roll-out of ART which occurred in the period between 2009 (the study year for Solarin and Black, 2013) and 2014, the latter being the survey year of this study. Given this roll-out, it is assumed that the disease burden would have improved which, if all other factors remained relatively constant, should translate into improved indicators in 2014. The indicator on the percentage of clients testing HIV-positive at their first antenatal care visit demonstrates a larger disease burden (from an HIV perspective) amongst antenatal care attenders in Johannesburg Region F compared to the Cape Town sub-districts. This is likely to influence all of the other indicators. The MMR is almost 6 times higher for HIV-positive women who give birth in public healthcare facilities compared to HIV negative women (Pattinson, 2012). However, even if this influence is taken into account, the indicators of the two Cape Town sub-districts, particularly indicators on both child and maternal mortality levels, seem to point towards slightly stronger supply-side performance.

⁵⁹ The data included here are averages for the full calendar year (indicated in the column headings). It was not possible to access the relevant data for the sub-district (Johannesburg Region F) and study year in which the Solarin and Black (2013) study was conducted and therefore include the district (Johannesburg Metropolitan) average. The sub-district data was, however, available for 2014. This serves as contextual comparison for the Cape Town Northern and Eastern sub-districts in 2014, the sub-districts in and year during which the survey was conducted.

⁶⁰ It seems maternal mortality and child fatalities were poorly captured for Metropolitan Johannesburg for this particular year. The very low maternal mortality level in 2009 seems to be a result of maternal mortality simply not being captured (11/12 months have zero rates, while the twelfth month had a rate of 35).

⁶¹ One of the district hospitals in the sample falls under this sub-district. It serves women both in this sub-district and in another sub-district.

5.8 Data

Below I provide information on the survey process and the data collected. A total of 221 women were interviewed shortly after giving birth in four public healthcare facilities in Metropolitan Cape Town.

Approach:

The interviews were conducted by five interviewers in English, isiXhosa and Afrikaans. The questionnaire, information and consent sheets were all available in English, isiXhosa and Afrikaans, although the majority of non-English speaking respondents preferred to use the English questionnaire, information and consent sheets. A small minority of respondents were non-South Africans who were less proficient in English. For these respondents, interviewers took more time to explain the survey in simpler language (English).

Similar to Solarin and Black (2013), the interviews were conducted at hospitals after birth rather than antenatal clinics. The interviews were conducted at hospitals after birth rather than at antenatal care clinics in order not to bias the quality of antenatal care received at the clinics.

The interviews took place by the bedside of women in the post-labour ward during the work week (Monday to Friday). Women in maternity obstetric units who have uncomplicated vaginal births can be discharged as soon as six hours after birth and for these women interviewers had only a small window of opportunity to find survey respondents. Women who had complicated births and, specifically, caesarean sections in the two district hospitals could be kept in the post-birth ward for as long as three days or even longer if there are post-birth complications.

Depending on the nature of responses, the interview duration was between 15 and 50 minutes. Interviews often took longer as interviewers waited patiently for mothers to finish their interactions with doctors, nurses and/or visitors. Some interviews were therefore subject to interruptions. According to feedback received from the interviewers, some respondents expressed their gratitude about the psycho-social support received from interviewers in the post-labour ward environment where doctors and nurses are mainly focused on their clinical duties.

The sample size for each facility was calculated based on the proportional contribution of the facility to total births amongst the four facilities in the year preceding the survey. Data for this calculation was obtained from the Western Cape Department of Health. The sample cannot be considered representative of the demography of the users of the facilities as convenience sampling was used

and mothers were interviewed according to their availability and willingness to participate in the study.

Survey structure:

I used the questionnaire originally developed by Solarin and Black (2013), which contains sections on socio-demographics, whether an antenatal care clinic was attended and when, reasons for attendance or non-attendance and experiences of care received. The results of this study and Solarin and Black's study are not directly comparable because of non-representative samples. The two surveys were conducted five years⁶² apart.

The original survey instrument contained sets of questions on detailed socio-demographic information, pregnancy discovery and antenatal care attendance (including possible reasons for late attendance, if relevant, and reasons why it is necessary to attend), the nature of the pregnancy and experiences of antenatal care received.

The section on socio-demographic questions from the original survey was expanded to include questions on asset ownership and more access to services questions to allow for the construction of an index to capture socio-economic status (SES). Where appropriate, language was also simplified. Questions on travel mode, travel time and cost of travel to the antenatal care clinic were included, while response options probing the potential role of fear of HIV testing in delayed antenatal care access were also included.

As with Solarin and Black's (2013) study, gestational age at first clinic attendance was self-reported and distinctions were made between first clinic visit and women's first booking (screening) visit to allow for identification of supply-side constraints if clinics were unable to book women at their first visit to the clinic. If women reported they had attended antenatal care at only 5 months or later, the survey identified them as "late attenders" and they were asked to complete a section of questions on reasons for late antenatal care attendance. Women who indicated they had not attended any antenatal care were also asked to respond to questions on possible reasons why antenatal care was not accessed.

The self-reported age at first clinic attendance was used to consider causes of late access and differences between late and early attenders in terms of the 20 weeks/5 months definition of antenatal care attendance (the same definition used by the National Department of Health). Women who reported attending antenatal care at 5 months or later were directed to a section of the

⁶² The survey which informs Solarin and Black (2013) was conducted during October 2009.

questionnaire designed to explore possible causes of late access. However, the data also allows for exploration of the socio-economic and risk factor differences between early and late attenders in terms of a 3-month definition of early antenatal care attendance.

Ethical clearance:

The study was approved by the ethics committee of Stellenbosch University (HS1021/2014) and the WCDOH provided permission to interview respondents in the four healthcare facilities. Permission was also obtained from the facility and labour ward managers at the four facilities. Written informed consent in the language of the respondent's choice (English, Afrikaans or isiXhosa) was obtained for all study respondents. The ethical clearance obtained limited the study to interviews with mothers aged 18 years and older. A small gift to the value of about R25 was provided as a sign of appreciation of respondents' time upon completion of the interview.

5.9 Methods

The collected data was analysed using univariate, bivariate and multivariate methods. Vaus (2002) recommends using a statistical method designed for bivariate analysis when considering the difference between any two variables. According to Aschengrau and Seage (2014), the chi-square test (χ^2). test is commonly used in to establish the statistical significance of the difference between two variables. Statistical differences in bivariate analysis between late and early attenders were therefore identified through the chi-square test (χ^2).

I used single equation linear models (Wooldridge, 2002) or linear probability models⁶³ (LPMs) for the multivariate analysis to explore the size and significance of the association between the prevalence of late access and various supply and demand factors. Since most of the included variables are binary, the coefficients can be interpreted as the marginal increase in the probability of lateness (or any other dependent variable) that is associated with specific explanatory variables. Due to the length of the chapter, the choice of variables for the linear probability models is discussed in Section 5.10.6 shortly before the results are described so the rationale is clear to the reader in the results section.

P-values <0.1 were treated as statistically significant in the bivariate and multivariate analyses⁶⁴.

⁶³ Although LPMs are prone to the problem of heteroskedasticity, this is likely to have a relatively small effect on the empirical results (Angrist and Pischke, 2008) and LPMs (relative to probits) allow for easy interpretation of coefficients.

⁶⁴ According to Raj (1998) significance tests at the 10% level are not uncommon in the field of development economics.

5.10 Results

5.10.1 Sample characteristics

A summary of the most salient socio-demographic characteristics of the sample is provided in Table 5.2 below. The majority of respondents (54.2%) were interviewed in English. More than half (54.3%) self-identified as black, 42.3% as coloured and the remaining respondents as white⁶⁵. A relatively high percentage (9.1%) of respondents indicated they were non-South African. There was also a high percentage of migrants (international and internal) present in the sample - 59.7% indicated they would describe another place than the Cape Town metropolitan area as their home, with 84.8% of this group being internal migrants. Of the group of internal migrants, 85.7% provided a place name in the Eastern Cape as the location of their “home home”.

⁶⁵ Although respondents were provided with an option of “Indian/Asian” as possible population group category, no respondents selected this option.

Table 5.2: Socio-demographic characteristics of total sample

Language of interview (n=219/221):	%
English	54.3
isiXhosa	33.3
Afrikaans	11.4
Other	0.9
Country of birth (n=220/221):	
South Africa	90.9
Other	9.1
Do you have another home (“home home”)? (n=221/221)	
No	40.3
Yes	59.7
Population group (n=221/221):	
Black	54.3
Coloured	42.5
White	3.2
Main work activity (n=221/221):	
Paid employment	38.9
Self-employed	4.5
Not employed	54.7
Schooling	1.8
Highest level of education (n=221/221):	
Primary school	3.6
Up to grade 11	49.3
Matric	39.8
Tertiary education (diploma/university)	5.9
Other	1.4
Relationship status (n=220/221)	
Single	46.3
Married	36.4
Living together	17.3

A large proportion of respondents indicated they were not employed (54.7%)⁶⁶, with 43.5% not having completed high school. More than half (53.7%) of the sample reported having a partner, with 36.4% being married and 17.3% living with a partner.

⁶⁶ Not employed respondents are respondents who selected the category “unemployed” in a question on main activity or work status. The question categories included paid employee, paid family worker, self-employed, employer, unpaid family worker and unemployed. The question did not distinguish between narrow and broad unemployment or adhere to an official definition of unemployment. Only one respondent selected “unpaid family worker” and this response has been grouped with “unemployed” which could now be interpreted as a “not employed” category. Given that the group of respondents who self-identified as “unemployed” is quite large (n=120 or 54.3%), this group could conceivably include individuals who are not unemployed according to a narrow or even broad definition of unemployed, but simply not economically active.

The mean age of respondents in the sample was 27.3 years (SD=5.8). About a quarter of the sample (24.9%) was aged between 18 and 22 years. The mean monthly household income of the sample was R3,024⁶⁷ and mean per capita household income was R701.

5.10.2 Pregnancy recognition, confirmation and care seeking

When asked how they first recognised pregnancy in terms of common pregnancy signs or symptoms⁶⁸, the majority of the sample indicated they skipped a menstrual cycle (72%) and/or felt nauseous (50.7%). The question included various other pregnancy symptoms⁶⁹. Under the response option of “other”, two (0.9%) women mentioned that they suspected they were pregnant when they felt the infant move.

Pregnancy confirmation:

The most popular pregnancy confirmation method was to buy and use a urine pregnancy testing kit (48.6%), followed by going to a primary healthcare clinic for a pregnancy test (39.6%) (Table 5.3). Following confirmation of pregnancy, 83.9% reported seeking care from the primary healthcare system, with 77.5% visiting an antenatal care clinic and 6.4% accessing another primary healthcare facility. The next biggest health provider category for care following pregnancy confirmation is a private doctor (GP), with 6.4% of respondents seeking care from this provider.

⁶⁷ The survey instrument asked about household income in seven income brackets, including a no income option. Mean household income was calculated by using a midpoint approach. See the original questionnaire for more information on the income questions (Appendix B to this chapter). Von Fintel (2006) found that the typical income bracket structure used in the South African General Household Surveys lends itself towards a midpoint approach without leading to coefficient bias. Relatively similar income brackets to those used in the General Household Survey were used in this chapter.

⁶⁸ Multiple mentions, spontaneous (unprompted) question. It is important to recognise that responses to this question cannot be considered a true reflection of knowledge of pregnancy symptoms at the start of the pregnancy as women’s knowledge on these symptoms may have increased during pregnancy and through their interaction with the healthcare system. Nevertheless, their responses may still provide some indication of pregnancy symptom knowledge.

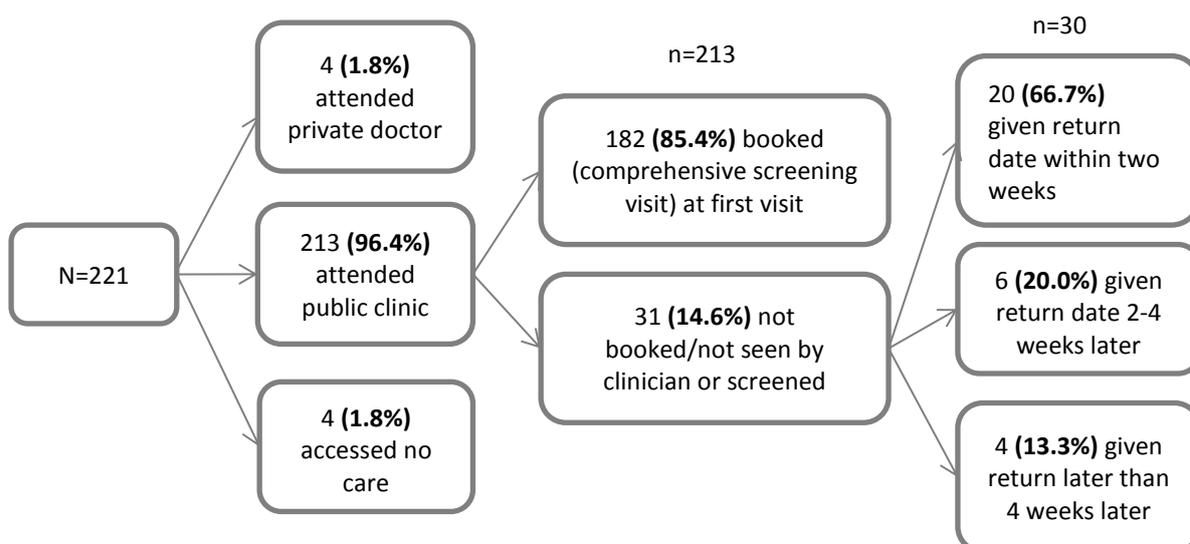
⁶⁹ Other pregnancy symptoms include sore or swollen breasts, feeling very tired, increased urination and craving certain foods.

Table 5.3: How pregnancy was confirmed

Pregnancy confirmation ⁷⁰ (n=220/221):	%
I bought a pregnancy test	48.6
I went to a clinic and was tested	39.6
I went to the clinic for something else and found out I was pregnant	3.6
I missed a few periods/menstrual cycles	2.3
I could feel the infant moving/kicking	0.5
I just knew I was pregnant	0.0

Similar to the high levels of antenatal care attendance reflected in both national and provincial administrative data for the public sector, a high proportion of women (96.4%) reported seeking care for their pregnancy at an antenatal clinic (Figure 5.5). The remainder either reported attending at a private doctor (1.8%) or not accessing any healthcare (1.8%) before birth. For those who attended antenatal care at a public clinic, the mean number of visits was 5.8 (SD=2.1) and the median number of visits was 6. A large proportion of women (86.4%) reported visiting the antenatal clinic four or more times during the pregnancy, the minimum number of times recommended by the WHO and implemented by the South African Department of Health through the Basic Antenatal Care (BANC) approach (Department of Health, 2007).

Figure 5.5: Antenatal care seeking cascade for all women



A large majority of women (85.4%) who attended antenatal care reported having their booking (full screening) visit the first time they visited the antenatal clinic (Figure 5.5). They were therefore seen

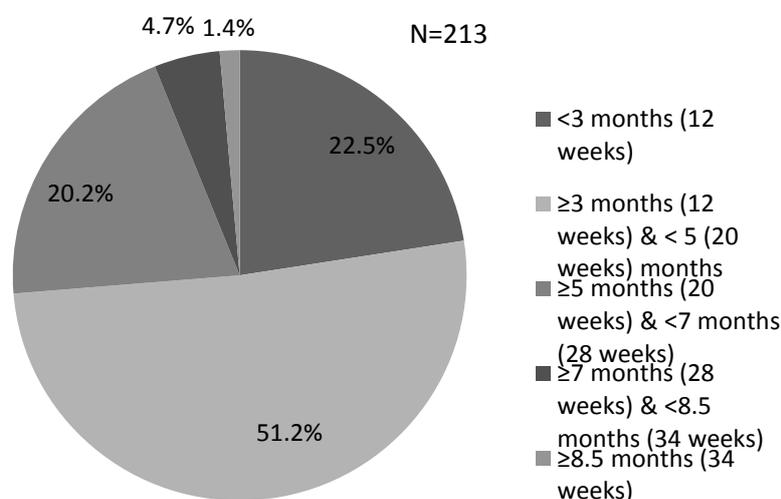
⁷⁰ 5.5% of women indicated some “other” method of pregnancy confirmation.

on the same day as their first visit by nursing staff and a number of basic health screenings were conducted. Of the 14.6% of women who did not have their booking visit at their first interaction with the antenatal care clinic and who provided data on whether they were asked to return later, 66.7% were given a return date within two weeks, 20% were given a return date 2-4 weeks after their visit and only 13.3% were given a return date for later than four weeks. Although the 14.6% of women who were not booked at their first visit is a substantially lower percentage of women than the finding by Solarin and Black (2013) of 49% of women not being booked at their first visit, this is still of much concern.

5.10.3 Timing of antenatal care attendance and self-reported causes of late access

Slightly more than a quarter (26.3%) of women who attended antenatal care self-reported attending antenatal care for the first time in their pregnancies at 5 months gestation or later (government definition of late attendance) (Figure 5.6). The majority (51.2%) visited the clinic for the first time between the first trimester (3 months) and 5 months, with only 22.5% of women reporting attendance during the first trimester (before 3 months).

Figure 5.6: Timing of antenatal care attendance



There was some variation in the 5-months/20 weeks late attendance rate amongst birth facilities – the highest late attendance rate was 33.0% and the lowest 18.0%. There were no statistically significant differences in timing of attendance rate by birth facility and also no strong pattern in terms of timing of attendance for the hospitals compared to the MOUs.

The 26.3% of women attending antenatal care for the first time at 5 months or later were asked a series of questions to establish causes of late attendance. The most frequently provided reason for late attendance was delayed recognition of pregnancy – 48.2% indicated they discovered they were pregnant quite late in their pregnancies (Table 5.4). The next most frequently cited reason was “other” which can be considered a catch-all explanation for a variety of free responses, followed by purposeful postponement (“I kept putting it off”) (19.6%) and not thinking it was necessary to go early (10.7%). Only 5.4% of women reported going late because the clinic asked them to return on another date. Some of the free responses provided under “other” overlap with the standard response options and if these are included with the standard response options in Table 5.4 below, “other” reasons were cited by only 17.9% of the sample, while late recognition of pregnancy increases to 51.8% and purposeful postponement to 23.2%.

Table 5.4: General reasons for late antenatal care attendance⁷¹

n=56/56	Given answers (%)	Include “other” reasons with standard response options if overlap (%)
I did not know I was pregnant until quite late	48.2	51.8
Other	30.4	17.9
I kept putting it off	19.6	23.2
I did not think it was necessary	10.7	10.7
I did not know I had to go any earlier than I did	5.4	5.4
I did not have time to go any earlier	5.4	10.7
I did not like the clinic	5.4	5.4
I did go before 20 weeks but was turned away/told to come back at a later date	5.4	5.4
I did not know where to go	0.0	0.0

Some of the remaining “other reasons” include reasons related to movement between provinces (“I was in the Eastern Cape and didn’t want to attend in the Eastern Cape”, 1.8%), fear of HIV testing (1.8%) and not having an identity document (3.6%). Not knowing where to go for antenatal care was not selected as an option by anyone, while none of the free responses directly indicated a lack of funds to travel to the clinic although one respondent noted experiencing transport problems under “other”.

⁷¹ This was a multiple response question and responses therefore do not add up to 100%, although the majority of respondents only selected one response.

Similar to the more general causes of late access, the most frequently provided response to the question on which personal factors could have enabled earlier attendance related to the timing of confirmation of pregnancy – 48.2% of respondents indicated they would have attended earlier if they had found they were pregnant sooner (Table 5.5 below). The second largest response item was “nothing would have made me go sooner”, with 26.8% of respondents selecting this response. For these women it is unlikely that supply-side interventions directly related to antenatal care would have made any difference to their health seeking behaviour. The response item “nothing would have made me go sooner”, together with the responses on encouragement “by those around me” (7.1%) and “if I had the support of those around me” could indicate quite a large need for psycho-social support during pregnancy amongst this sample of women.

Table 5.5: Personal factors that could have enabled earlier attendance⁷²

n=56/56	Given answers (%)
If I found out sooner that I was pregnant	48.2
Nothing would have made me go sooner	26.8
Other	12.7
If I thought there might be a problem	12.5
If I thought it was necessary	10.7
If I was told or knew I had to go	10.7
If I had time to go sooner	7.1
If I did not feel well	7.1
If I was encouraged by those around me	7.1
If I felt I had the support of those around me	7.1

When asked about clinic factors that could have enabled earlier attendance, the highest number of responses to a single item was to “nothing about the clinic would have made me go sooner” (42.3%) (Table 5.6). The next biggest response item was “other” with 26.9% of late attendees selecting this option and providing free responses in the follow-up question. Only one of the free responses indicated a clinic factor may have influenced health seeking behaviour (“no clinic in the area”). The remaining free responses all related to personal factors such as late discovery of pregnancy (19.2%), denial or non-acceptance of pregnancy (1.9%), having limited time available due to work time constraints (1.9%) and ascribing late attendance to non-clinic factors (1.9%, “clinic had nothing to do

⁷² This was a multiple response question and responses therefore do not add up to 100%, although the majority of respondents only selected one response.

with my choice”). If these responses are re-assigned to the given categories with which they overlap, the percentage of respondents who indicated that the clinic had no impact on their late attendance behaviour increases to more than two thirds of late attendees (67.3%), while “other” decreases to 1.9% of late attendees.

The most prominent supply-side theme that emerges from the responses relate to the ability of the clinic to see more patients during a day and, more specifically, include waiting times (13.5%), opening and closing times (9.6%) and having attended earlier but then being asked to return or come back, which indicates rationing (11.5%). Two respondents indicated that a lack of funds for transport to the clinic may have contributed to their late attendance (3.8%), while one respondent (1.9%) indicated she may have attended antenatal care earlier in her pregnancy if she did not have been tested for HIV at the clinic.

Table 5.6: Clinic factors that could have enabled earlier attendance

Clinic factors that could have enabled earlier attendance (n=52/56)	Given answers (%)	Include “other” reasons with standard response options if overlap (%)
Nothing about the clinic would have made me go sooner	42.3	67.3
Other	26.9	1.9
If I did not have to wait too long at the clinic	13.5	13.5
If the clinic staff were more friendly	11.5	11.5
I went earlier but I was turned away/told to come back	11.5	11.5
If the clinic opening and closing times were better	9.6	9.6
If the clinic was closer to me	7.7	7.7
If I liked the clinic	3.8	3.8
If I had money to go to the clinic	3.8	3.8
If I did not have to be tested for HIV at the clinic	1.9	1.9

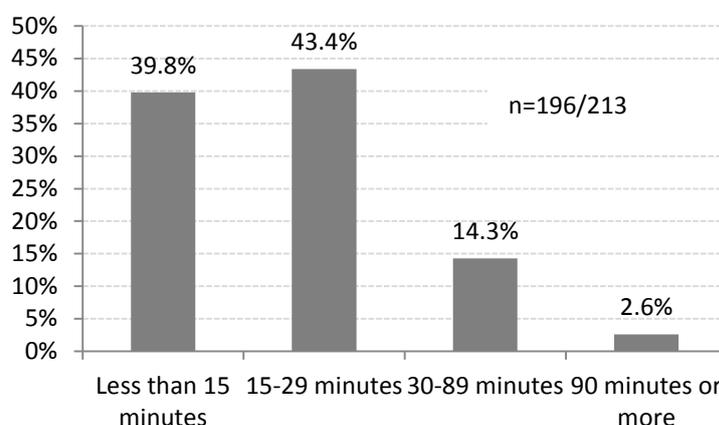
5.10.4 Supply-side factors and timing of attendance

In this section, the potential supply-side barriers which may be associated with late attendance are explored. Some of these can be considered slightly more objective measures than the self-identified

causes of late access discussed previously and therefore provide a validation method relative to the causes discussed in the previous section. Three categories of access barriers are considered: physical access, affordability as captured by travel costs and quality or user acceptability.

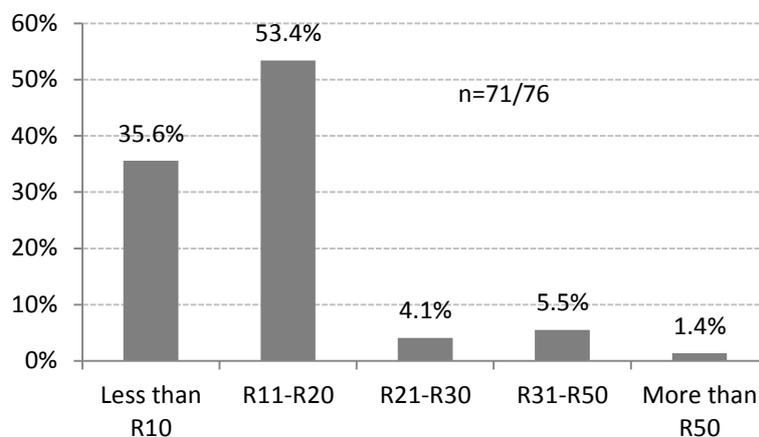
The two most common modes of transport to the clinic were walking (55.2%) and use of public transport (36.3%). Of 196 respondents who provided travel time to the clinic, 83.2% reported travelling for less than 30 minutes per one-way trip (see Figure 5.7 below). Only 16.9% reported traveling for 30 minutes or more to reach the clinic, with 2.6% reporting travel times of 90 minutes or more. Travel time does not seem to be a prohibitive factor in terms of timing of access to antenatal care.

Figure 5.7: Travel time to clinic for those who attended antenatal care



Respondents who indicated they used public transport to travel to the clinic were asked about the cost of a one-away trip to the clinic. The majority of respondents (52.3%) indicated that such a trip costs between R11 and R20 (Figure 5.8), which does not seem prohibitively high. Depending on the type of pregnancy (high vs. low-risk) and also how advanced the pregnancy was at the time of booking, more than four visits may have been required during the pregnancy. More than four visits could translate into relatively high travel cost expenditure given the low mean monthly per capita income of R701 in the sample.

Figure 5.8: Cost of one-way trip to clinic for those who attended antenatal care and used public transport to the clinic



I did not find any significant association between late attendance and longer travel time or travel cost. Clinic proximity and associated travel expenditure therefore do not appear to be a major access barrier in terms of timing of antenatal care attendance amongst this sample of women.

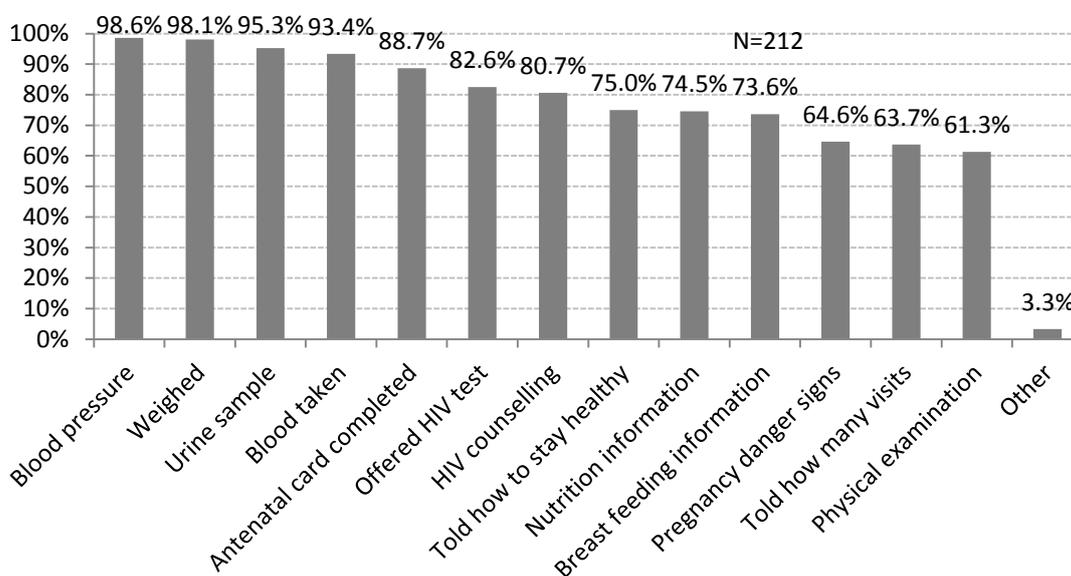
The survey included a question on whether certain services that are typical of a first full antenatal screening visit were provided to respondents at their booking visit. The most prominent (and generic) services were included in the question. This question allows for a basic assessment of the comprehensiveness and therefore basic quality of services provided to the patient.

At the booking visit, patients are assessed and classified as having a low or high risk pregnancy. This classification relies on factors related to their obstetric history (e.g. birth weight of previous infants, complications experienced during previous pregnancies), the current pregnancy (e.g. aged <16 years, aged 37 years or older at conception, vaginal bleeding) and their general medical condition (e.g. having been diagnosed and treated for diabetes mellitus, substance abuse, having been diagnosed and treated for TB, etc.) (Western Cape Government: Health, 2015c). A detailed questionnaire is completed to gather this information. Based on the classification, the patient may be required to have more visits than the minimum number prescribed by the Basic Antenatal Care (BANC) approach or be referred to a higher level of care, i.e. antenatal care at a hospital. A copy of the most recent Western Cape Department of Health clinic checklist “Booking and follow-up of HIV negative patient”, together with the standard blue “Maternity case record” booklet indicates that all the services included in the questionnaire, drawing on the original survey instrument of Solarin and Black (2013), indeed have to be provided at the booking visit.

The question on services received at the booking visit may be subject to large recall bias given elapsed time between the booking visit and giving birth and the specificity of services asked about. Nevertheless, responses may still be indicative of whether certain services were provided and also whether these services were discussed and/or explained to respondents (respondents who were informed about why certain services were provided and what the results of screening tests were may have better recall).

Figure 5.9, below, summarises response rates for whether the various service elements were provided. In general, response levels seem high, with a slight decrease in responses on questions related to HIV services (whether an HIV test was offered and HIV counselling). However, the service elements that relate to the provision and sharing of information with antenatal attendees (e.g. nutrition information, breast feeding, pregnancy danger signs) had quite low response rates compared to the more “visible” service elements (e.g. blood pressure measurement, being weighed, a urine sample being taken). This may reflect the difficulty of recalling information rather than reflecting on services that were provided at the visit.

Figure 5.9: Services that were provided by clinic at booking (first screening) visit



I tested for significance of association between groups of women who did not report certain services being offered by the clinic and late attendance (Table 5.7, below) with the hypothesis that if a particular clinic or the clinic system overall delivers poorer quality services or is perceived to deliver poor quality services, women may delay seeking care from this clinic. The only service element for which a significant association between reporting that this service was not offered and late antenatal care access was found, is HIV counselling. Women who accessed antenatal care at 5 months or later were less likely to report HIV counselling not being offered at their antenatal care

clinic during the booking visit. This association is in the opposite direction than anticipated if considered purely as an indication of service quality (comprehensiveness of service offering). However, if fear of HIV testing and status confirmation plays a role in the timing of antenatal care attendance as previously identified (e.g. Amnesty International, 2014), women who delayed antenatal care attendance due to fears related to HIV may be reluctant to talk about this topic which could bias their responses. No statistically significant association was found for late attendance and whether an HIV test was offered or not.

Table 5.7: Services not offered by clinic at booking visit (early vs. late attenders)

*** p<0.01, ** p<0.05, * p<0.1

Booking service not offered	Total (%)	Early (<5 months)	Late (≥5 months)	p-value (early vs. late)
Blood pressure	3 (1.4%)	1 (0.6%)	2 (3.6%)	0.111
Weighed	4 (1.9%)	3 (1.9%)	1 (1.8%)	0.948
Urine sample	10 (4.7%)	7 (4.5%)	3 (5.4%)	0.792
Blood taken	14 (6.6%)	11 (7.1%)	3 (5.4%)	0.661
Antenatal card completed	24 (11.3%)	16 (10.3%)	8 (14.3%)	0.414
HIV test	37 (17.5%)	25 (16.03%)	12 (21.43%)	0.361
HIV counselling	41 (19.3%)	36 (23.1%)	5 (8.9%)	0.021**
Told how to stay healthy	53 (25.0%)	41 (26.3%)	12 (21.4%)	0.472
Nutrition information	54 (24.5%)	38 (24.4%)	16 (28.6%)	0.535
Breast feeding information	56 (26.4%)	39 (25.0%)	17 (30.4%)	0.435
Pregnancy danger signs	75 (35.4%)	51 (32.7%)	24 (42.9%)	0.172
Told how many visits	77 (36.3%)	60 (38.5%)	17 (30.4%)	0.279
Physical examination	82 (36.7%)	59 (37.8%)	23 (41.1%)	0.668

On a collective level, whether certain clinic services were provided or not could serve as an indication of perceptions of overall health service quality, which may influence health seeking behaviour if women do not experience services as adding any value to their pregnancy and/or health. To create a collective measure for the quality of health service provision, I estimated a health service index through multiple correspondence analysis (MCA) of the various individual health services reported in Figure 5.8 and Table 5.6. MCA is a statistical method which identifies patterns in the underlying relationships between variables to identify a common element (typically the first dimension), which can be used to estimate a single score (Greenacres and Blasius, 2006). I used an MCA rather than principal component analysis (PCA) approach for this index because Booyesen et al. (2008), in their construction of an asset index for seven African countries, argued that MCA is a more appropriate approach for the construction of an index which uses data from categorical rather than continuous variables. (More information on these indices can be found in the next section which

includes a description of the asset index used as proxy for socio-economic status in exploring correlates of late access.)

Following the calculation of a health service index score for each respondent, it was attempted to identify groups of scores. Attempts to cluster the health service index scores into quintiles or quartiles indicated that the distribution of scores best lends itself to clustering in three groups (terciles). No significant associations were found between the terciles of the health services index and both definitions of late access (≥ 5 months and ≥ 3 months). The results are not reported here.

Overall there were very low levels of reported dissatisfaction with easy-to-measure clinic quality indicators which were included in the survey. There were low total numbers of respondents who expressed some form of dissatisfaction or even neutrality (i.e. the absence of approval) with regards to various positive statements about how clinic services were experienced (Table 5.8, below). For example, only 4.7% of respondents expressed some form of dissatisfaction or neutrality to the statement “I was generally satisfied with the care I received”. There was some uncertainty about how to interpret the significant association between respondents reporting they had a problem getting to the clinic and late access ($P < 0.05$). The question was intended as a *physical access* question but respondents may have interpreted this as relating also to other access barriers they experienced. The negative or neutral responses to this statement could reflect transport or other physical access constraints or respondents could have interpreted this as relating to their own delayed antenatal care access, irrespective of whether this related to supply-side or demand-side constraints.

Table 5.8: Reported weak performance on clinic quality indicators by early vs. late antenatal care attendance

*** p<0.01, ** p<0.05, * p<0.1

Sub-sample who disagreed, strongly disagreed or were neutral/indifferent about these statements:	Total (%) (N=213)	Early (<5 months) (N=157)	Late (≥5 months) (N=56)	p-value (early vs. late)
Staff at the clinic were very friendly	17 (8.0%)	13 (8.3%)	17 (8.0%)	0.778
Staff at the clinic behaved how health providers should behave	17 (8.1%)	12 (7.7%)	5 (8.9%)	0.780
Staff at the clinic knew what they were talking about	4 (1.9%)	3 (1.9%)	1 (1.8%)	0.944
The clinic looked clean	11 (5.2%)	8 (5.1%)	3 (5.5%)	0.925
The amount of time I had to wait to be seen was reasonable	33 (15.6%)	25 (16.0%)	8 (14.6%)	0.795
I had a problem getting to the clinic	9 (4.25%)	4 (2.6%)	5 (8.9%)	0.043**
Staff provided me with all the information I needed	11 (5.3%)	7 (4.6%)	4 (7.3%)	0.443
The staff communicated very well with me	17 (8.1%)	11 (7.1%)	6 (10.7%)	0.394
I was generally satisfied with the care I received	10 (4.7%)	7 (4.5%)	3 (5.4%)	0.800

5.10.5 Demand-side factors and timing of attendance

In this section, the nature of associations between late antenatal care attendance and a variety of demand-side factors is explored. The demand-side factors include the socio-economic, demographic, health and risk factor data collected in the survey. These groups of data describe the patient or possible health seekers' lives and therefore point towards demand-side factors that may influence health seeking, also the timing of antenatal care attendance. It is important to understand the relative importance of these factors for knowing whether remedial policy should be directed towards the supply- or demand-side.

Characteristics of women who accessed antenatal care late (government definition, ≥5 months):

Table 5.9 (below) contains a summary of some of the main characteristics of women who accessed antenatal care late according to the official definition of late access, namely access at or after 5 months gestation. Given the longer period of time before women are considered to be late in accessing antenatal care, this is also referred to this as the broader definition of late access.

A higher percentage of women who experienced a second or later pregnancy accessed antenatal care at 5 months gestation or later compared to women who experienced a first pregnancy (76.8% vs 23.2%) ($P < 0.05$). This is in contrast to the finding by Solarin and Black (2013) in metropolitan Johannesburg where a first pregnancy was significantly associated with late access of antenatal care.

The percentage of coloured respondents was slightly higher amongst late attenders compared to early attenders (46.4% vs. 40.8%) but the race group variable was not significant in terms of late access. While a large proportion of late attenders reported being single (as opposed to married or living with a partner) relative to early attenders (50.0% vs. 45.5%), this difference is also not significant.

A higher proportion of late attenders reported not having completed high school compared to early attenders (69.6% vs. 48.4%, $P < 0.01$). There was no significant association between employment status relative to the timing of antenatal care attendance, and the distribution of the two categories of work status (employed vs. not employed) was quite similar amongst the late attenders compared to early attenders. There were high levels of unemployment amongst both groups, with 58.6% of early attenders indicating they were not employed compared to 57.1% for late attenders.

There was no statistically significant difference between the two groups in terms of formality of housing. This result is not reported in Table 5.9 as the same variable was also included in the asset index measure (discussed later). A planned pregnancy relative to two categories of unplanned pregnancy (“Unplanned and unhappy that I am pregnant”, “Unplanned but happy that I am pregnant”) is not significantly associated with late access. The prevalence of any alcohol consumption during pregnancy⁷³, a potential indicator of risky behaviour and life choices, was higher amongst late attenders (30.3%) compared to early attenders (15.9%) ($P < 0.05$).

⁷³ A binary variable created by combining the responses to two questions: 1) During your pregnancy, how often did you drink any alcoholic beverages; 2) During your pregnancy, did you ever drink alcoholic beverages until you passed out? Indicating any type of alcohol consumption relative to a combination of the two questions was captured as a 1 while no alcohol consumption relative to a combination of the two questions was captured as 0.

Table 5.9: Socio-economic characteristics and health behaviour by government definitions of early (<5 months) vs. late (≥5 months) antenatal care attendance

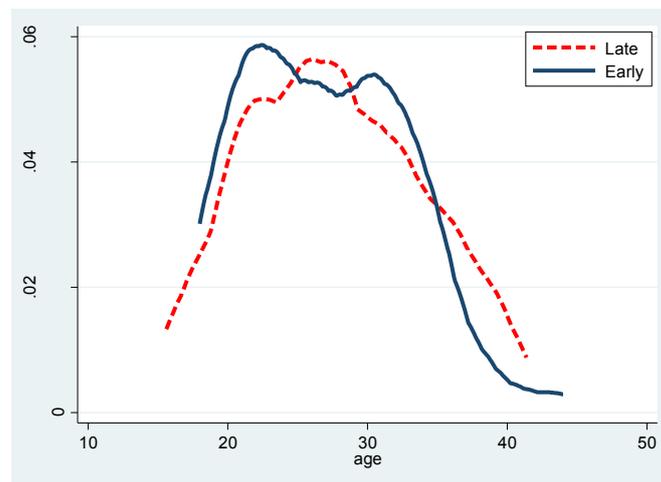
*** p<0.01, ** p<0.05, * p<0.1

Characteristics	Total (%)	Early (<5 months) (%)	Late (≥5 months)	p-value (early vs. late)
Gravida/pregnancy number (including current) (n=213/213)				
1 st	74 (34.7%)	61 (38.9%)	13 (23.2%)	0.035**
Second or later	139 (65.3%)	96 (61.2%)	43 (76.8%)	
Population group/race (n=213/213)				
Black	117 (54.9%)	89 (56.7%)	28 (50.0%)	0.669
Coloured	90 (42.3%)	64 (40.8%)	26 (46.4%)	
Other	6 (2.8%)	4 (2.6%)	2 (3.6%)	
Marital status (n=212/213)				
Single	99 (46.7%)	71 (45.5%)	28 (50.0%)	0.369
Living together	76 (35.9%)	60 (38.5%)	16 (28.6%)	
Married	37 (17.5%)	25 (16.0%)	12 (21.4%)	
Education (n=213/213)				
Less than high school/matric	115 (54.0%)	76 (48.4%)	39 (69.6%)	0.006***
Completed high school/matric	98 (46.0%)	81 (51.6%)	17 (30.4%)	
Employment status (n=213/213)				
Employed (paid employment and self-employed)	89 (41.8%)	65 (41.4%)	24 (42.9%)	0.850
Not employed	124 (58.2%)	92 (58.6%)	32 (57.1%)	
Household income (n=186/213)				
Household income ≤R3,200 per month	122 (65.6%)	92 (65.7%)	30 (65.2%)	0.951
Household income >R3,200 per month	64 (34.3%)	48 (34.3%)	16 (34.8%)	
Position in asset index (asset wealth) (n=213/213)				
Top 60%	128 (60.1%)	101 (64.3%)	27 (48.2%)	0.034**
Bottom 40%	85 (39.9%)	56 (35.7%)	29 (51.8%)	
Planned pregnancy (n=212/213)				
Yes	47 (22.2%)	37 (23.7%)	10 (17.9%)	0.243
No, but happy that I am pregnant	107 (50.5%)	81 (51.9%)	26 (46.4%)	
No and unhappy that I am pregnant	58 (27.4%)	38 (24.4%)	20 (35.7%)	
Consumed any alcohol during pregnancy (n=213/213)				
Yes	42 (19.7%)	25 (15.9%)	17 (30.3%)	0.020**
No	171 (80.3%)	132 (4.1%)	39 (69.6%)	

There were no easily observable differences in the distribution of ages of late vs. early attenders (government definition, ≥5 months) (Figure 5.10). The mean age of respondents who attended antenatal care early is 27.0 years compared to the mean age of respondents who attended antenatal

care late: 27.6. This difference was not statistically significant. Various age cut-off points were used to create binary age groupings, but no statistically significant differences were found between early and late attenders.

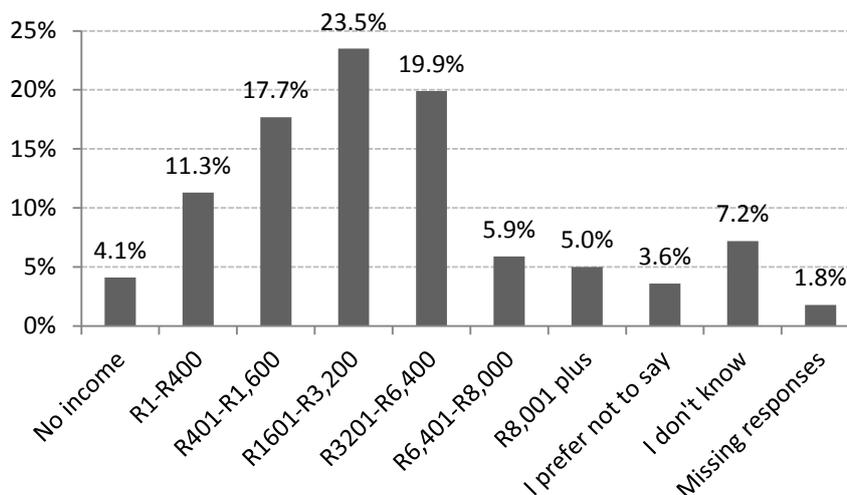
Figure 5.10: Distribution of age of booked by early (<5 months) vs. late antenatal care attendance (≥5 months)



The survey collected data on both personal and household income in seven income brackets. Household income generally is a more reliable income measure than personal income. Where income is reported on it thus refers to household income. A number of respondents did not provide any information on household income (5.4%)⁷⁴, chose to indicate they receive no household income⁷⁵ (4.1%) or indicated that they did not know how much income their household receives (7.2%) (see Figure 5.11). If these respondents are excluded, income data is available for only 83.3% of total respondents and 83.1% of respondents who attended antenatal care.

⁷⁴ Prefer not to say and missing responses

⁷⁵ In some cases, respondents in this group indicated they received some type of government grant or provided information on source of income which showed (contrary to their response on the question on household income) their household indeed receives income.

Figure 5.11: Percentage of observations (total sample, N=221) by household income response category

Despite the limited number of responses on household income data, I still report summary measures as an approximate indication of income amongst late and early attenders. The mean per capita household income of early attenders (government definition, ≥ 5 months) was R701 per month, compared to a mean per capita household income of R664 for late attenders per month. These differences are not statistically significant.

In order to maximise the number of respondents used in correlations and regression analysis, I elected to use a measure of socio-economics status, an asset index approach, to include as large a number of respondents as possible. This approach was pioneered by Filmer and Pritchett (2001) when they were the first authors to construct an asset index as proxy for household wealth or socio-economic status and to demonstrate that it produced similar results as using household expenditure data. They did so while examining the relationship between household wealth and school enrolment in India. Since then, large numbers of authors using mainly principal component analysis (PCA) but more recently also multiple correspondence analysis (MCA), have estimated and used asset indices as proxy for household wealth or socio-economic status (see, for example, the short review of studies using asset indices in Booyesen et al., 2008). Asset indices have also been used in South African health studies to explore the relationship between household wealth or socio-economic status and different health outcomes or utilisation measures, for example, child health (Nkonki et al., 2011) and access to antenatal care services (Wabiri et al., 2013).

Data on the ownership of 14 different durable assets by the household were collected and were used in the estimation of an asset index to serve as proxy for household socio-economic status. A further two variables that capture access to services were included in the index. This includes whether the respondent reported living in what could be considered a formal housing structure and

whether the respondent reported access to running water in the home. A total of 16 variables were therefore used in the construction of the index. Table 5.10 (below) provides a summary of the frequency at which ownership of the assets included in the survey were reported. If compared to ownership levels of the same assets by a weighted sample of women who indicated they are pregnant and reside in urban areas in the most recent wave (2012) of NIDS, ownership of durable assets in the study sample of women who recently gave birth in metropolitan Cape Town appears quite similar to national figures.

Table 5.10: Ownership of durable assets from survey vs. ownership of selected same assets from NIDS (2012)

Asset or service category	Percentage of total antenatal care attendance sample at public sector clinics (N=213):	Percentage of pregnant women (ages 18-44) in urban areas in South Africa (NIDS, 2012):
Cell phone	96.2%	92.8%
(Electric) stove	95.7%	88.0%
Television	90.6%	81.8%
Fridge	81.5%	77.8%
Hifi/stereo	73.3%	42.4%
Running water in home	70.3%	65.8%
Stays in formal house/building	66.7%	73.2%
Radio	66.4%	67.8%
Microwave	64.3%	64.6%
Washing machine	46.7%	51.8%
Satellite	28.9%	36.2%
Computer	21.6%	18.5%
Car	21.3%	32.6%
Paraffin stove	17.1%	16.5%

As recommended by Booysen et al. (2008) for the construction of an asset index using categorical variables (such as the available asset and access to services variables), an MCA approach was used to estimate the asset index. The factor loadings of the individual variables are reported in Table 5.A1 in Appendix A to this chapter. The ownership of assets I hypothesised to have a positive association with higher income levels contribute positively to asset index scores (Table 5.A1). There is only one asset, paraffin stove, for which ownership negatively contributes to asset index scores. The first dimension which results from the MCA explains 89.5% of inertia and it is this dimension which is used to estimate the asset index. Asset index scores range from -2.7 to 1.7.

Falling into the bottom 40% of the asset index score distribution, i.e. falling into the poorest 40% of this socio-economic status measure, is significantly associated with late antenatal care attendance

(Table 5.9). The same pattern of significance did not hold for a broader grouping of asset index poverty, namely falling into the bottom 60% of the asset distribution. No significant patterns of association between timing of access and household income were found.

Characteristics of women who accessed antenatal care late (alternative definition, ≥ 3 months):

Table 5.11, below, contains a summary of the main characteristics of women who accessed antenatal care late relative to the alternative definition of late access, i.e. presenting at an antenatal care clinic for the first time after the first trimester (3 months) of pregnancy. Given that more women are likely to only attend at 3 months gestation or later, I also refer to this definition as the broad definition of late access.

As with the 5 months late access definition, there were a large proportion of women who experienced a second or later pregnancy amongst late attenders compared to early attenders (66.7% vs. 60.4%). There were proportionally more black women amongst late relative to early attenders (56.3% vs. 50.0%). Neither of the reported differences was statistically significant.

Table 5.11: Socio-economic characteristics and health behaviour by alternative definitions of early (<3 months) vs. late (≥ 3 months) antenatal care attendance

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Characteristics	Total (%)	Early (<3 months) (%)	Late (≥ 3 months)	p-value (early vs. late)
Gravida/pregnancy number (including current) (n=213/213)				
1st	74 (34.7%)	19 (39.6%)	55 (33.3%)	0.423
Second or more	139 (65.3%)	29 (60.4%)	110 (66.7%)	
Population group/race (n=213/213)				
Black	117 (54.9%)	24 (50.0%)	93 (56.4%)	0.650
Coloured	90 (42.3%)	23 (47.9%)	67 (40.6%)	
Other	6 (2.8%)	1 (2.1%)	5 (3.0%)	
Marital status (n=212/213)				
Single	99 (46.7%)	26 (54.2%)	73 (44.5%)	0.286
Living together	76 (35.9%)	17 (35.4%)	59 (36.0%)	
Married	37 (17.5%)	5 (10.4%)	32 (19.5%)	
Education (n=213/213)				
Less than high school/matric	115 (54.0%)	23 (47.9%)	92 (55.8%)	0.337
Completed high school/matric	98 (46.0%)	25 (52.1%)	73 (44.2%)	
Shack in informal settlement	19 (8.9%)	5 (10.4%)	14 (8.5%)	

Employment status (n=213/213)				
Employed (paid employment and self-employed)	89 (41.8%)	25 (52.1%)	64 (38.8%)	0.100*
Not employed/schooling	124 (58.2%)	23 (47.9%)	101 (61.2%)	
Household income (n=186/213)				
Household income ≤R3,200 per month	122 (65.6%)	31 (73.8%)	91 (63.2%)	0.203
Household income >R3,200 per month	64 (34.3%)	11 (26.2%)	53 (36.8%)	
Position in asset index (asset wealth) (n=213/213)				
Top 60%	128 (60.1%)	33 (68.7%)	95 (57.6%)	0.164
Bottom 40%	85 (39.9%)	15 (31.3%)	70 (42.4%)	
Planned pregnancy (n=212/213)				
Yes	47 (22.2%)	12 (25.0%)	35 (21.3%)	0.030**
No, but happy that I am pregnant	107 (50.5%)	30 (62.5%)	77 (47.0%)	
No and unhappy that I am pregnant	58 (27.4%)	6 (12.5%)	52 (31.7%)	
Consumed any alcohol during pregnancy (n=213/213)				
Yes	42 (19.7%)	12 (25.0%)	30 (18.2%)	0.296
No	171 (80.3%)	36 (75.0%)	135 (81.8%)	

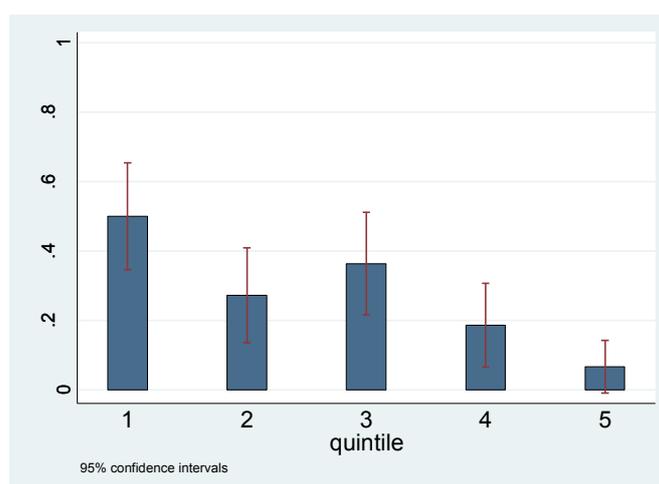
The majority of women (55.8%) who attended antenatal care reported not having completed high school. This was not the case for early attenders, where 52.1% indicated they had completed high school, but the differences are not statistically significant ($P=0.337$).

Women who were not employed were significantly more likely to only attend antenatal care at or after 3 months (61.2% vs. 47.9%, $P=0.1$). This relationship is in the opposite direction than the one found by Solarin and Black (2013), where the employed were more likely to only access antenatal care at 5 months or later, reflecting some of the supply-side constraints found in the same study.

There was a lower proportion of single women amongst late attenders relative to early attenders (44.5% vs. 54.2%, $P=0.286$). Differences in late attendance for whether the pregnancy was planned and whether women welcomed it in their lives were statistically significant. There was a larger proportion of women who indicated they had an unplanned pregnancy about which they were unhappy amongst late attenders compared to early attenders (31.7% vs. 12.5%, $P<0.05$). There was no statistically significant difference in reported alcohol consumption amongst late relative to early attenders.

A higher percentage of late attenders reside in households with monthly incomes of less than R3,200 (36.8% vs. 26.2%, $P=0.203$). Furthermore, a slightly higher proportion of late attenders compared to early attenders fell into the bottom two quintiles of the asset index (42.4% vs. 41.3%, $P=0.164$) but this difference was not significant. However, the difference between the higher proportion of late attenders who fell into the bottom three quintiles of the asset index compared to early attenders (81.4% vs. 71.4%) was statistically significant ($P<0.1$, not reported in Table 5.10). It is possible that this result may therefore be driven by women who fell into the third quintile of the asset index distribution. Figure 5.12, below, reveals that a smaller proportion of women who accessed antenatal care late according to the broader definition fell into the second compared to the first and third quintiles. The addition of a third quintile could help generate significant results in terms of socio-economic status difference in timing of attendance.

Figure 5.12: Proportion of women who accessed antenatal care late (≥ 3 months) by asset index quintile



Late attendance using the broader definition, as with the narrower definition, is thus also associated with indicators of socio-economic vulnerability, although less severe vulnerability as there is only a significant association with falling into the bottom three quintiles and not the bottom two quintiles of the asset index distribution.

Characteristics of women by desirability of pregnancy in their lives:

Given the statistical significance of differences in the desirability of pregnancy amongst late compared to early attenders, the characteristics of women by this category were explored. The analysis distinguished between women who reported having an unplanned pregnancy about which they were unhappy and women who indicated they had either a planned pregnancy (for which it is assumed they were happy about) or an unplanned pregnancy about which they reported being happy. These latter two categories were grouped together as the statistical significance of results in

the bivariate analysis seemed to have been driven by the desirability of the pregnancy. Table 5.12, below, summarises results for some of the main characteristics of these two groups of women. For ease of reference, for the remainder of the discussion these two groups of pregnancies are distinguished by referring to them as unwanted and wanted pregnancies.

Table 5.12: Socio-economic and health characteristics/behaviour of respondents by unwanted pregnancies vs. wanted pregnancies

*** p<0.01, ** p<0.05, * p<0.1

Behaviour and characteristics	Total (%)	Unwanted pregnancies (%)	Wanted pregnancies (%)	p-value (early vs. late)
Government definition of late attendance (>5 months) (n=212/213)				
Late	56 (26.4%)	20 (34.5%)	36 (23.4%)	0.102
Early	156 (73.4%)	38 (65.2%)	118 (76.6%)	
Alternative definition of late attendance (>3 months) (n=212/213)				
Late	164 (77.4%)	52 (89.7%)	112 (72.73%)	0.009***
Early	48 (22.6%)	6 (10.3%)	42 (27.3%)	
Gravida/pregnancy number (including current) (n=212/213)				
1st	74 (34.9%)	20 (34.5%)	54 (35.1%)	0.937
Second or more	138 (65.1%)	38 (65.5%)	100 (64.9%)	
Population group/race (n=213/213)				
Black	116 (54.7%)	37 (63.8%)	79 (51.3%)	0.123
Coloured	90 (42.5%)	21 (36.2%)	69 (44.8%)	
Other	6 (2.8%)	0 (0.0%)	6 (3.9%)	
Relationship status (n=211/213)				
Single	99 (46.9%)	36 (62.1%)	63 (41.2%)	0.001***
Living together	75 (35.6%)	9 (15.5%)	66 (43.1%)	
Married	37 (17.5%)	13 (22.4%)	24 (15.7%)	
Education (n=212/213)				
Less than high school/matric	115 (54.3%)	37 (63.8%)	78 (50.7%)	0.087*
Completed high school/matric	97 (45.8%)	21 (36.2%)	76 (49.4%)	
Housing/accommodation (n=212/213)				
Formal house/townhouse/flat/shared house/room in house	141 (66.5%)	33 (56.9%)	108 (70.1%)	0.073*
House/flat in backyard	52 (24.5%)	16 (27.6%)	36 (23.4%)	
Shack in informal settlement	19 (9.0%)	9 (15.5%)	10 (6.5%)	

Employment status (n=212/213)				
Employed (paid employment and self-employed)	96 (42.6%)	25 (41.0%)	71 (44.7%)	0.623
Not employed/schooling	124 (56.4%)	36 (59.0%)	88 (55.4%)	
Household income (n=185/213)				
Household income ≤R3,200 per month	122 (66.0%)	40 (81.6%)	54 (39.7%)	0.007***
Household income >R3,200 per month	63 (34.0%)	9 (18.4%)	82 (60.3%)	
Position in asset index (asset wealth) (n=212/213)				
Top 60%	127 (59.9%)	27 (46.6%)	100 (65.9%)	0.015**
Bottom 40%	85 (40.1%)	31 (53.5%)	54 (35.1%)	
Consumed any alcohol during pregnancy (n=212/213)				
Yes	42 (19.8%)	12 (20.7%)	30 (19.5%)	0.844
No	170 (80.2%)	46 (79.3%)	124 (80.5%)	

A higher percentage of women with unwanted pregnancies compared to women with wanted pregnancies attended antenatal care late according to both the narrow and the broad definition. However, these differences were only statistically significant for late attendance according to the broad definition (≥ 3 months) ($P < 0.01$). The proportion of first pregnancies amongst women with unwanted and wanted pregnancies was very similar (34.5% vs. 35.1%). There was a much higher proportion of single women amongst women with unwanted pregnancies compared to wanted pregnancies (62.1% vs. 41.2%, $P < 0.01$).

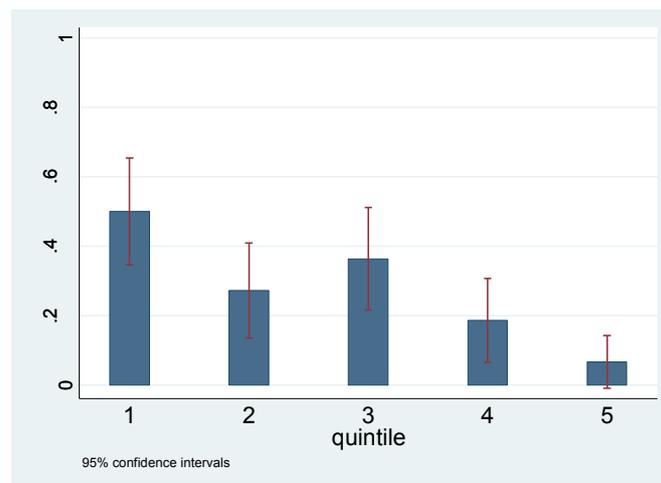
There was a higher proportion of black women amongst women with unwanted pregnancies compared to women with welcome pregnancies (63.8% vs. 51.3%). A larger percentage of women with unwanted pregnancies reported not having completed high school compared to women with wanted pregnancies (63.8% vs. 50.7%, $P < 0.1$), while women with unwanted pregnancies were also more likely to indicate they lived in a shack in an informal settlement and less likely to report residing in a formal dwelling structure ($P < 0.1$). Women with unwanted pregnancies were also slightly more likely to report being not employed relative to women with wanted pregnancies (59.0% vs. 55.4%).

The proportion of women in both groups (unwanted vs. wanted pregnancies) who reported any alcohol consumption was very similar (20.7% vs 19.5%).

Women with unwanted pregnancies were more likely to be part of households which had monthly incomes of lower than R3,200 ($P < 0.01$). Similar to differences in the socio-economic status and the broader definition of late attendance, the results for differences in the distribution of women with “unwanted” pregnancies by socio-economic status seems to be driven by women falling into the first

and third quintiles of the asset index distribution (see Figure 5.13 below). There was a higher proportion of women with unwanted pregnancies compared to women with wanted pregnancies who fell into both simple definitions of relative asset poverty, i.e. falling into either the bottom 40% of the asset index distribution ($P < 0.05$, see Table 5.12) or falling into the bottom 60% of the asset index distribution ($P < 0.01$, not reported in Table 5.12).

Figure 5.13: Proportion of respondents who indicated they have an unplanned pregnancy about which they are unhappy (unwanted pregnancies) by asset index quintile



There was a higher proportion of women with unwanted pregnancies for almost all indicators of socio-economic vulnerability, including lower education levels, lower household income and relative asset poverty. Alcohol consumption, and gravida, which could influence perceptions on the necessity of early antenatal care attendance given healthy previous pregnancies, seemed less important for this group of women.

5.10.6 Regression analysis results

The analysis was limited by the small sample size which restricted me to variables that were available for most respondents. The analysis therefore risks missing out on identifying the importance of association between late access and excluded factors⁷⁶ that only apply to sub-groups of women. The results have to be interpreted with caution as they show association and do not allow for the inference of causal relationships between late access and supply and demand factors.

Two categories of binary dependent variables were used: late access in terms of the narrow (5 month) definition and late access in terms of the broad (3 month) definition. For both definitions of late access six different specifications were run. In specification 1, I controlled for the most important demand-side (socio-demographic) factors, two factors describing the pregnancy (gravida and whether the pregnancy was planned) and alcohol consumption to potentially pick up risky health behaviour choices. All of the included explanatory variables, except age, are binary. In specifications 2 and 3 two binary supply-side measures were added: firstly, whether the respondent reported having good access to antenatal care services and, secondly, whether the respondent expressed some form of satisfaction with the way antenatal care health services were provided. In specifications 3-6 the same approach was used keeping almost all variables the same, but replacing the planned pregnancy variable with a variable to control for desirability of pregnancy. These results are reported in Tables 5.14 and 5.15.

Given the importance of “desirability” of pregnancy as identified through the descriptive analysis, a separate set of regressions for desirability⁷⁷ was run using binary variable to control for unwanted pregnancies as the dependent variable. The same approach as described above was used, except the explanatory variable on whether the pregnancy was planned was omitted due to the collinear relationship with desirability of pregnancy (Table 5.16).

Under the assumption that these might be different types of pregnancies, separate regressions were run for women by unwanted and wanted pregnancies (Tables 5.A2 and 5.A4) and women by first or second and later pregnancies (Tables 5.A3 and 5.A5 in Appendix 5.A), for both definitions of late access (narrow and broad).

Table 5.13 summarises the expected signs of coefficients and the rationale for the expected signs.

⁷⁶ E.g. experiencing complications during a previous pregnancy: only a subsample of women has been pregnant more than once and could therefore report on whether or not they experienced complications during previous pregnancies.

⁷⁷ Unwelcome pregnancies, as defined earlier, take on a value of 1 and welcome pregnancies take on a value of 0.

Table 5.13: Variables included in regression analysis – expected sign and rationale

Variable	Expected sign	Rationale
Black (African)	Positive	Control for race as socio-demographic factor in South African context as it may capture dimensions of vulnerability and exclusion not fully captured by the socio-economic status measures included in regression.
Age	Negative	Younger women may have few peers that have had children and can therefore share knowledge and information about pregnancy health seeking. Younger women may be more reluctant to seek care due to stigma around teenage pregnancy.
Grant recipient	Unsure	Being a grant recipient is an indication of socio-economic vulnerability which could be positively associated with delayed or late access. However, the grant itself may put women in a better position to fund opportunity costs associated with care seeking, thereby improving early access. Depending on nature of the grant, may be highly correlated with already having a young child (previous pregnancy).
Completed high school	Negative	Education provides access to knowledge and information which is likely to influence health seeking behaviour.
Employed	Unsure	Employed women may delay antenatal care access due to time constraints, restricted opening hours of clinics or long waiting times. Dependent on the formality of employment, employed women may also not be willing to forego income associated with a clinic visit. On the other hand, women who are not employed may have limited income to fund travel and other associated opportunity costs.
Asset index controlling for socio-economic status	Positive	Falling lower down in the index indicates relative poverty and vulnerability.
Partner	Negative	A partner (here indicating co-habitation with a partner or being married as opposed to being single) may provide financial and emotional support during and after pregnancy, thereby lightening the childcare burden.
Gravida: first or second (or later) pregnancies	Unsure	Women who experience a first pregnancy may struggle to recognise and interpret pregnancy symptoms. These women may also not yet understand the importance of antenatal care. On the other hand, women who experience a second or later pregnancy, who had healthy first pregnancies, may underestimate the importance of early antenatal care.
Planned pregnancy	Negative	A planned pregnancy is an anticipated pregnancy - women who plan to get pregnant may be on the lookout for pregnancy symptoms (be in a “pregnancy state of mind”) and therefore be in a position to take earlier healthcare action.

Unwanted pregnancy (desirability of pregnancy)	Positive	An unwanted pregnancy is likely to be associated with denial and potentially purposeful postponement of care seeking.
Alcohol consumption during pregnancy	Positive	This can be considered an indication of risky behaviour and life choices which may also be associated with risk(ier) health behaviour.
Satisfied with antenatal care (clinic) services received	Negative	Women who perceive or experience government primary health care services as satisfactory may be more willing to use these services.
Good access to antenatal care (clinic) services	Negative	Women who do not report access barriers to public antenatal care services may be more likely to visit these services and earlier during pregnancy.

For all sets of regressions, the specific asset poverty definition was varied according to the significance of results obtained from the cross-tabulations with the dependent variable (reported in an earlier section). The definition that provided statistically significant results was used. For some regressions I included an asset poverty dummy variable to indicate whether respondents fell into the bottom 40% of the asset index distribution, while for other regressions I used an asset poverty dummy variable to indicate whether respondents fell into the bottom 60% of the asset index distribution.

Late access in terms of the narrow definition (≥ 5 months) is significantly associated with not having completed high school ($P < 0.1$) and reported alcohol consumption during pregnancy ($P < 0.1$) (Table 5.14). These associations remain consistent, except for regression specifications 2 and 5. In specification 2, where *planned* pregnancy (rather than *unwanted* pregnancy) is controlled for, the significance of alcohol consumption disappears when a control is included for having good access to antenatal care⁷⁸. In specification (5) (where *unwanted* pregnancy rather than *planned* pregnancy was controlled for) the high school completion variable becomes insignificant with the addition of good access variable.

⁷⁸ The only specification for which “good access” is itself significant.

Table 5.14: Regression results for late access (narrow definition, ≥5 months) to antenatal care

Variables	(1) late	(2) late	(3) late	(4) late	(5) late	(6) late
Black	-0.0283 (0.0661)	-0.0270 (0.0661)	-0.0293 (0.0665)	-0.0373 (0.0661)	-0.0349 (0.0662)	-0.0370 (0.0666)
Grant	-0.0182 (0.0764)	-0.0257 (0.0764)	-0.0214 (0.0772)	-0.0187 (0.0761)	-0.0245 (0.0761)	-0.0200 (0.0769)
High school	-0.115* (0.0655)	-0.111* (0.0654)	-0.116* (0.0666)	-0.110* (0.0655)	-0.107 (0.0654)	-0.111* (0.0668)
Employed	0.0445 (0.0655)	0.0268 (0.0663)	0.0246 (0.0666)	0.0439 (0.0654)	0.0278 (0.0662)	0.0257 (0.0666)
Asset poor (bottom 2 quintiles)	0.114 (0.0693)	0.0979 (0.0698)	0.0935 (0.0704)	0.105 (0.0695)	0.0913 (0.0700)	0.0876 (0.0706)
Partner	-0.0333 (0.0664)	-0.0278 (0.0663)	-0.0267 (0.0669)	-0.0282 (0.0660)	-0.0262 (0.0659)	-0.0256 (0.0666)
First preg	-0.116 (0.0798)	-0.126 (0.0798)	-0.129 (0.0806)	-0.115 (0.0796)	-0.124 (0.0797)	-0.127 (0.0805)
Planned preg	-0.0475 (0.0747)	-0.0539 (0.0750)	-0.0574 (0.0756)			
Alcohol	0.141* (0.0807)	0.137* (0.0806)	0.132 (0.0812)	0.142* (0.0804)	0.139* (0.0803)	0.136* (0.0810)
Age	-0.00398 (0.00637)	-0.00464 (0.00640)	-0.00452 (0.00643)	-0.00403 (0.00634)	-0.00464 (0.00639)	-0.00456 (0.00642)
Good access		0.263* (0.152)	0.249 (0.160)		0.238 (0.153)	0.229 (0.160)
Satisfied			0.0473 (0.151)			0.0306 (0.151)
Unwanted preg				0.0823 (0.0699)	0.0689 (0.0703)	0.0669 (0.0708)
Constant	0.424** (0.192)	0.447** (0.193)	0.451** (0.194)	0.396** (0.193)	0.420** (0.194)	0.424** (0.195)
Observations	210	209	208	210	209	208
R-squared	0.085	0.097	0.098	0.090	0.100	0.099

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

The only variable that is significantly associated with the broad definition of late access (≥ 3 months) is unwanted pregnancy, which is included in specifications 4-6 (Table 5.15). Reporting an unwanted pregnancy as opposed to a wanted pregnancy increases the probability of late access (broad definition) by almost 17 percentage points ($P < 0.05$). When planned pregnancy rather than *unwanted* pregnancy is controlled for, as was done in specifications 1-3, none of the included variables are significantly associated with the broad definition of late access. The absence of more statistically significant coefficients could indicate that women who accessed antenatal care late in terms of the broad definition are a more heterogeneous group than women who were late in terms of the narrow definition⁷⁹.

⁷⁹ It could also be that more of the reasons for going slightly late (≥ 3 months) versus very late (≥ 5 months) are based on unobservable personality characteristics, whereas very late postponement is based on observable external circumstances and observable and measured traits.

Table 5.15: Regression results for late access (broad definition, ≥ 3 months) to antenatal care

Variables	(1) late2	(2) late2	(3) late2	(4) late2	(5) late2	(6) late2
Black	0.0145 (0.0629)	0.0108 (0.0633)	0.0145 (0.0635)	-0.000619 (0.0621)	-0.00382 (0.0625)	0.000203 (0.0628)
Grant	-0.0199 (0.0745)	-0.0199 (0.0749)	-0.0258 (0.0756)	-0.0226 (0.0733)	-0.0207 (0.0737)	-0.0266 (0.0744)
High school	-0.0273 (0.0655)	-0.0274 (0.0658)	-0.0399 (0.0673)	-0.0223 (0.0646)	-0.0235 (0.0650)	-0.0336 (0.0665)
Employed	-0.0779 (0.0637)	-0.0831 (0.0647)	-0.0813 (0.0649)	-0.0788 (0.0629)	-0.0805 (0.0639)	-0.0784 (0.0641)
Asset poor	0.0779 (0.0667)	0.0802 (0.0671)	0.0752 (0.0681)	0.0481 (0.0670)	0.0509 (0.0674)	0.0483 (0.0684)
Partner	0.0671 (0.0646)	0.0682 (0.0649)	0.0616 (0.0652)	0.0853 (0.0635)	0.0845 (0.0638)	0.0771 (0.0642)
First preg	-0.0497 (0.0763)	-0.0508 (0.0767)	-0.0604 (0.0772)	-0.0453 (0.0753)	-0.0451 (0.0757)	-0.0534 (0.0763)
Planned preg	-0.0518 (0.0727)	-0.0590 (0.0735)	-0.0647 (0.0739)			
Alcohol	-0.0771 (0.0779)	-0.0759 (0.0782)	-0.0790 (0.0785)	-0.0743 (0.0766)	-0.0726 (0.0770)	-0.0744 (0.0773)
Age	-0.00142 (0.00621)	-0.00202 (0.00627)	-0.00187 (0.00628)	-0.00149 (0.00612)	-0.00193 (0.00619)	-0.00182 (0.00620)
Good access		0.0525 (0.148)	0.000453 (0.155)		-0.00154 (0.147)	-0.0463 (0.154)
Satisfied			0.171 (0.149)			0.147 (0.146)
Unwanted preg				0.167** (0.0682)	0.168** (0.0690)	0.166** (0.0692)
Constant	0.813*** (0.192)	0.830*** (0.194)	0.835*** (0.195)	0.771*** (0.190)	0.782*** (0.192)	0.786*** (0.193)
Observations	210	209	208	210	209	208
R-squared	0.040	0.042	0.049	0.066	0.067	0.072

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Given the significance of unwanted pregnancy in terms of the broad definition (≥ 3 months) of late access, the characteristics of women who reported unwanted pregnancies were further explored (Table 5.16). In specifications 1-3, the factors that emerge as having a consistently significant association with reporting a pregnancy as unwanted are the absence of a partner (negative coefficient on the partner binary variable) and falling into the lowest 60% of the asset index distribution. The significance of association of the presence/absence of a partner may signal concerns about caring for and raising a child with only one parent present. Earlier exploration of the asset index variable in relation to unwanted pregnancies showed that the majority of women in this group fall into the first and third quintiles of the asset index distribution, therefore using a narrower asset poverty definition (falling into bottom two quintiles) does not yield significant results (specifications 4-6). Women who report not having a partner are almost 16 percentage points more likely ($P < 0.01$) than women with partners to report having an unwanted pregnancy, while falling into the bottom three quintiles of the asset index distribution increases the probability of an unwanted pregnancy by between 18 and 19 percentage points ($P < 0.01$).

Table 5.16: Regression results for unwanted pregnancies

VARIABLES	(1) Unwanted preg	(2) Unwanted preg	(3) Unwanted preg	(4) Unwanted preg	(5) Unwanted preg	(6) Unwanted preg
Black	0.0717 (0.0642)	0.0687 (0.0642)	0.0685 (0.0646)	0.0807 (0.0666)	0.0820 (0.0666)	0.0806 (0.0669)
Grant	0.0412 (0.0760)	0.0330 (0.0759)	0.0342 (0.0768)	0.0561 (0.0768)	0.0491 (0.0768)	0.0513 (0.0776)
High school	-0.0272 (0.0670)	-0.0200 (0.0669)	-0.0276 (0.0686)	-0.0675 (0.0661)	-0.0634 (0.0660)	-0.0733 (0.0672)
Partner	-0.160** (0.0649)	-0.156** (0.0647)	-0.158** (0.0653)	-0.159** (0.0658)	-0.154** (0.0657)	-0.157** (0.0662)
First preg	-0.0263 (0.0781)	-0.0333 (0.0779)	-0.0378 (0.0787)	-0.0153 (0.0805)	-0.0252 (0.0805)	-0.0324 (0.0812)
Alcohol	0.00627 (0.0795)	0.00529 (0.0792)	0.00236 (0.0798)	0.0239 (0.0813)	0.0203 (0.0811)	0.0146 (0.0817)
Age	-0.000991 (0.00634)	-0.00171 (0.00637)	-0.00164 (0.00640)	-0.00222 (0.00641)	-0.00286 (0.00645)	-0.00269 (0.00647)
Good access		0.275* (0.150)	0.251 (0.158)		0.260* (0.153)	0.224 (0.161)
Satisfied			0.0771 (0.151)			0.115 (0.152)
Employed	-0.000544 (0.0652)	-0.0184 (0.0658)	-0.0192 (0.0662)	-0.00617 (0.0661)	-0.0236 (0.0669)	-0.0250 (0.0672)
Asset poor	0.191*** (0.0682)	0.187*** (0.0681)	0.182*** (0.0693)			
Asset poor (bottom 2 quintiles)				0.100 (0.0699)	0.0843 (0.0704)	0.0799 (0.0710)
Constant	0.240 (0.197)	0.257 (0.197)	0.263 (0.198)	0.353* (0.194)	0.375* (0.195)	0.379* (0.195)
Observations	210	209	208	210	209	208
R-squared	0.100	0.114	0.114	0.075	0.087	0.088

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

For both definitions of late access, factors associated with late access were differentiated by desirability of pregnancy (unwanted vs. wanted pregnancies) and also by first and second (or later) pregnancies, as there may be large differences in health behaviour in terms of these different states. When access to antenatal care (narrow definition, ≥ 5 months) is differentiated by by unwanted and wanted pregnancies (Table 5.A2), only one variable is significantly associated with late access. The consumption of any alcohol during pregnancy increases the probability of late access by 17.6-19.3 percentage points across specifications 5-8 for women with *unwanted* pregnancies⁸⁰($P<0.1$). The absence of more precisely estimated coefficients in these regressions could be a consequence of the smaller sample size which results when regressions are separately estimated. Next late access is

⁸⁰ Where the binary variable takes on a value of 0.

considered by first and second (or later) pregnancies. For women who had a first pregnancy, the absence of a partner is associated with an increase in the probability of late access (narrow definition, ≥ 5 months) of about 21 percentage points ($P < 0.1$, specifications 1-4 in Table 5.A3).

When the factors potentially associated with the second definition of late access (narrow, ≥ 3 months) are differentiated by unwanted or wanted pregnancies, none of the variables are statistically significant (Table 5.A4). However, when late access is explored by first and second (or later) pregnancies, three variables seem to matter. Employment status is significantly associated with this definition of late access and only for women who had first pregnancies ($P < 0.01$). The probability of accessing antenatal care late for women who had first pregnancies increases by about 30 percentage points if women reported any type of employment (specifications 1-4, Table 5.A5). The variable is insignificant for women who had second or later pregnancies, but for women who had second or later pregnancies having a partner is significantly associated with late access ($P < 0.05$, specifications 5-8 in Table 5.A5). This is in contrast to the earlier results on the absence of a partner being significantly associated with the narrow definition of late access for first pregnancies. For women who had second or later pregnancies, reporting an unwanted pregnancy increases the probability of late access by about 16 percentage points ($P < 0.1$, specifications 5-8, Table 5.A5).

I also explored the associations between most of the variables contained in Table 5.A6 and timing of antenatal care access by pregnancy months. For this dependent variable, the coefficients have to be interpreted in terms of their contribution to delayed care by month. Reporting an unwanted pregnancy leads to a delay of about 0.6 of a month in accessing antenatal care (specifications 1-4, Table 5.A6, $P > 0.01$). This is quite a large delay if interpreted relative to the constant ($P < 0.1$) of about 3.5 months, not taking the association with any other factors into account. The narrower definition of asset poverty, falling into the bottom two quintiles of the asset distribution index, is also significantly ($P < 0.05$ in specification 1, $P < 0.1$ in specification 2) associated with delay by month. Falling into this part of the asset distribution is associated with a delay of 0.5 months.

5.11 Overview of findings: making sense of results

This study found an average late antenatal care attendance rate of 26.3% (narrow definition, ≥ 5 months). Given the trend in increases in early attendance rates in the Cape Town Metropolitan health district (Figure 5.4), an early attendance rate of 50-60% and therefore a late attendance rate of 40-50% were expected. The average late attendance rate is therefore much lower than anticipated. At first glance this finding appears counter-intuitive because one would expect over-reporting of early access in administrative data, given the considerable policy focus on increasing early access.

However, there are a number of possible explanations for the lower than anticipated late access rate. Firstly, the sample is not statistically representative of women attending antenatal care in the sampled sub-districts. This means that in interpreting the results less weight should be given to the exact late access rate. Rather, the results serve as a first indication of likely causes of late access among women in Metropolitan Cape Town. Also, given that the analysis relied on self-reported timing, it is possible that recall and social desirability bias may have affected women's responses. Social desirability bias would be the case if women knew it was recommended to access antenatal care earlier than they did and wanted to provide an answer in accordance with general societal or health system expectations. It is also possible that due to the roll-out of basic antenatal care services to more clinics managed by the City of Cape Town and a greater focus on early access promotion, the early antenatal care attendance rate increased at a greater rate in the survey year compared to earlier years. Data from the DHIS confirm large increases in early attendance rates by the middle of 2014, with early attendance rates in excess of 60%. Lastly, women younger than 18 years were not interviewed, given the study's ethical clearance conditions. It is possible that late access rates are higher amongst teenage mothers, which is not captured in the survey data.

No real evidence was found of supply-side factors being closely associated with late antenatal care attendance. When women who attended late (narrow definition, ≥ 5 months) were asked about clinic-related factors that could have enabled earlier attendance, 42.3% directly indicated that nothing about the clinic would have changed their behaviour. This figure increases to 67.3% if "other" responses which overlapped with this type of response are reassigned. A relatively small percentage of women were turned away at their first visit (14.6%) and asked to return at a later date – only 10 (33.3%) of those turned away were given a return date later than 2 weeks after their first visit. While this is lower than the levels reported in the Solarin and Black (2013) study, this practice should not be occurring in the health sector and is contrary to Department of Health Policies.

User acceptability does not seem to be a major constraint to early antenatal care access. The only booking service significantly related to late access (broad definition, ≥ 3 months) was not being offered HIV counselling at the booking appointment, but this was in the opposite direction than expected, with women who accessed antenatal care early more likely to report the absence of this service. Studies show that women may delay antenatal care access if they feel antenatal care services offer little value and do not warrant the opportunity cost of engaging with the health system (Abrahams, Jewkes and Mvo, 2011; Downe et al., 2009; Finlayson and Downe, 2013; Mayer and Harrison, 2003). This counter-intuitive finding on HIV counselling could, however, indicate that women who book late according to the broad definition may be more hesitant to report honestly on HIV services and topics. This would be the case if fear of HIV testing played a role in late antenatal care access or if some women feared a follow-on question about their HIV status. Although the survey included questions on fear of HIV testing influencing early access, only one respondent directly indicated that this was a reason for late access. This was thus not identified as an important factor influencing late access. It is possible, however, that a structured interview does not lend itself to clear and open responses on HIV, which is a heavily stigmatised disease.

More than half of women (51.8%) who delayed antenatal care access to 5 months or later of pregnancy reported not knowing that they were pregnant as the main reason for this delay. This could indicate that many women have difficulty recognising the signs of pregnancy from a knowledge perspective, or that there may be some biological or behavioural reason why it is difficult to identify pregnancy.

The survey did not ask questions about contraceptive usage and methods. If women struggled to link pregnancy symptoms to the realisation of pregnancy, it is possible that they did not anticipate pregnancy due to contraceptive usage. The low likelihood of pregnancy anticipation is confirmed by the high level of unplanned pregnancies amongst women in the sample. Various studies report relatively high levels of unplanned pregnancies amongst sub-samples of South African women (Macphail et al., 2007; Schwartz et al., 2012; Solarin and Black, 2013). The last available national data on contraceptive prevalence rates amongst women aged 15-49 indicate that 59.9% women reported any type of modern contraception usage in 2003/04 (United Nations, 2014). According to an estimate for contraception usage in 2010, this may have increased to 63.7% (Alkema et al., 2013). The high level of unplanned pregnancies in the study, however, raises questions about how widespread contraception usage was amongst women in the study sample and potentially also the efficacy of chosen contraceptives.

Even if contraception is used, some may mask pregnancy symptoms. Depo Provera, an injectable contraceptive and one of the most frequently used form of contraception in South Africa, could lead to irregular menstrual cycles or the absence of menstruation (amenorrhoea). It has been reported as making early pregnancy identification difficult, thereby contributing to delayed attendance (Abrahams, Jewkes and Mvo, 2001). High levels of unplanned pregnancies amongst women receiving ART therapy have also raised questions about ART contributing to injectable contraceptive failure (Schwartz et al., 2012).

Even if women are using contraceptives and do not anticipate pregnancy, easy access to pregnancy testing may aid in early identification of pregnancy. Usage of urine pregnancy tests or better access to pregnancy tests in South Africa has been shown to be significantly associated with a lower gestational age at presentation for antenatal care (Jeffery et al., 2000; Morroni and Moodley, 2006) and abortion services (Morroni and Moodley, 2006).

The survey also provides tentative evidence of denial of pregnancy and possible purposeful postponement of antenatal care attendance. This theme is common to the literature on late antenatal care attendance (Downe et al., 2009; Haddrill et al., 2014). A relatively large proportion (26.8%) of late attenders (narrow definition, ≥ 5 months) indicated that nothing would have made them attend earlier, with a further 14.2% indicating they would have gone earlier if they were encouraged by those around them or felt they had the support of others around them.

It should be clear from the above discussion that while there is some evidence of supply-side constraints, particularly availability of healthcare services, the main reasons for late access seem to be demand-side related, with relatively strong associations between late attendance, indicators of socio-economic vulnerability and behavioural factors. This finding is in contrast to that of the study of Solarin and Black (2013), where a large part of delayed access seemed due to supply-side constraints at antenatal care clinics in inner-city Johannesburg. A much larger proportion of women in the Solarin and Black (2013) study compared to my study were not seen by clinics at their first visit, causing long delays in accessing comprehensive screening services. The earlier comparison of health system indicators in this study sub-districts to indicators from the Solarin and Black (2013) study sub-district supports the idea of a stronger health supply-side in the two sub-districts included in this study.

The correlates of late attendance identified through this study cannot be interpreted as causal factors, but controlling for a number of factors through regression analysis at least allows for an understanding of the relative importance and direction of possible influence of these correlates in

late care seeking. Here I briefly comment on the nature of the relationship between late access and two variables in the regression results: unwanted pregnancies and the reported presence of a partner. The importance of unwanted pregnancies, which includes both a planning component and women's emotional experience of the pregnancy, as a correlate of late access, is emphasised by the regression results. When women who reported unwanted pregnancies are considered separately in regression analysis, the absence of a partner (negative coefficient) emerges as a significant correlate of regarding the pregnancy as unwanted. This may point towards poor access to or sub-optimal behavioural choices around contraception.

Limitations:

It was not possible to analyse health outcomes (both for the mother and their infants) with the data collected, given the limits of the ethical clearance obtained. A further limitation imposed by the limits of the ethical clearance relates to late access among women younger than 18 years. Anecdotal evidence indicates that late access levels may be much higher amongst this group of women. I, however, did not have permission to administer the survey to women younger than 18 years. It would also have been interesting to compare the actual timing of antenatal care access with self-reported timing. This comparison is important for understanding the accuracy of self-reported timing for use in future studies.

5.12 Conclusion

The chapter shows that there is not only one factor contributing to delayed antenatal care attendance in the sample of women in the Western Cape. Rather, the women in this study's sample display large levels of heterogeneity in terms of the likely factors that affected delayed antenatal care attendance. In the data and analysis, late antenatal care attendance is a multi-faceted behavioural choice or health outcome that most probably requires multi-pronged policy responses to address the problem. To conclude, some policy considerations and options are highlighted.

Supply-side constraints did not seem to be the major driver of late antenatal care attendance. However, this study identified that even in what appears to be a more functional supply-side context than in the Solarin and Black (2013) study, there are still some women who report being turned away at clinics for antenatal booking visits. Clinics may have developed individualised operating systems where they only accept a certain number of women for booking visits per day or only on certain days or at certain times. This is not in keeping with Department of Health policy and shows a deviation from clearly developed health protocol. Women being turned away can lead to further delays in attendance in a context where early identification of HIV and initiation of ARVs can make a

large difference to both the mother and infant's health. It is not clear what the immediate solution to this practice is, but it is possible to at least speculate about its causes. It is possible that this type of antenatal care is delivered in settings with weaker management practices or with lower levels of accountability relative to provincial and national healthcare policy. This will require further investigation and remedial action.

I identified delayed recognition of pregnancy as the major cause of late antenatal care access and also found a high level of unplanned pregnancies. This raises questions about the scale of contraceptive access, usage and knowledge amongst sampled women. Appropriate and efficacious contraception has the potential to globally prevent large numbers of maternal deaths (Ahmed et al., 2012). Contraception has been identified as a key mechanism to prevent maternal deaths in South Africa (Moodley, 2014). The large focus on HIV prevention in South Africa is likely to have diminished the policy and delivery focus on contraception for pregnancy prevention. The data and analysis may indicate that contraception usage was not sufficiently widespread amongst women in the sample and a better understanding of contraception access and choices is required.

Delayed recognition of pregnancy also raises questions about access to pregnancy confirmation methods. Making urine pregnancy tests widely available in the public sector could have a large impact on women's ability to detect pregnancy as early as possible, empowering them to make choices about antenatal care access or termination of pregnancy in a timely manner. Termination of pregnancy is currently offered up to 12 weeks of pregnancy by a nurse and up to 20 weeks of pregnancy by a doctor (Western Cape Government: Health, 2015e). A large proportion of women in this study may have missed these cut-off periods or if they discovered their pregnancy before these cut-offs, would not have had much time to consider their options.

Lastly, the analysis demonstrates the importance of information and possibly also incentives to allow women to better discount the benefits of antenatal care into current decision-making and behaviour. If South Africa is to increase early antenatal care attendance rates to optimal levels for HIV identification and treatment and the treatment of other conditions contributing to maternal mortality, experimentation with a variety of policy levers is required. This includes providing more and better information and financial or other incentives to women for better decision-making.

Implications for future research:

Future research should include questions on contraception methods used at the time of conception. A further line of research could focus on possible health reasons why women were unable to recognise pregnancy. If possible, more objective health indicators such as weight and height (to

calculate BMI) and even blood pressure readings should be collected. The research also needs to be expanded to mothers younger than 18 years as this is a vulnerable group of women.

Appendix A to Chapter 5

Table 5.A1: Variables used in estimation of asset index and weights/factor loadings obtained from MCA

Variables	Asset ownership	Weights
Formal house	Yes	0.860
	No	-1.618
Water in home	Yes	0.935
	No	-1.974
Fridge	Yes	0.587
	No	-2.441
Hifi stereo	Yes	0.681
	No	-1.885
Radio	Yes	-0.617
	No	-1.162
Television	Yes	0.277
	No	2.739
Satellite dish	Yes	1.703
	No	-0.648
Computer	Yes	1.484
	No	-0.417
Cell phone	Yes	0.062
	No	-1.661
Stove	Yes	0.107
	No	-2.879
Paraffin stove	Yes	-0.741
	No	0.134
Microwave oven	Yes	1.024
	No	-1.763
Washing machine	Yes	-1.299
	No	1.081
Car	Yes	2.006
	No	-0.564

Table 5.A2: Regression results for late access (narrow definition) by unwanted vs. wanted pregnancies

Variables	Unwanted pregnancies				Wanted pregnancies			
	(1) late	(2) late	(3) late	(4) late	(5) late	(6) late	(7) late	(8) late
Black	-0.115 (0.157)	-0.192 (0.162)	-0.168 (0.169)	-0.234 (0.171)	0.0122 (0.0744)	0.0193 (0.0753)	0.000522 (0.0748)	0.00713 (0.0758)
Grant	0.0624 (0.168)	0.0558 (0.169)	0.0525 (0.158)	0.0897 (0.160)	-0.0389 (0.0899)	-0.0477 (0.0925)	-0.0510 (0.0903)	-0.0569 (0.0928)
High school	-0.236 (0.157)	-0.157 (0.162)	-0.178 (0.159)	-0.129 (0.161)	-0.102 (0.0768)	-0.116 (0.0785)	-0.0953 (0.0744)	-0.108 (0.0757)
Employed	-0.0769 (0.154)	-0.206 (0.171)	-0.0706 (0.153)	-0.176 (0.165)	0.0672 (0.0748)	0.0599 (0.0754)	0.0662 (0.0743)	0.0599 (0.0749)
Asset poor	-0.0385 (0.196)	0.0976 (0.210)			0.0442 (0.0762)	0.0256 (0.0775)		
Partner	-0.163 (0.147)	-0.143 (0.146)	-0.118 (0.157)	-0.104 (0.155)	0.00211 (0.0776)	0.00209 (0.0787)	-0.00863 (0.0779)	-0.00943 (0.0794)
First pregnancy	-0.232 (0.187)	-0.241 (0.185)	-0.188 (0.199)	-0.167 (0.200)	-0.104 (0.0874)	-0.122 (0.0884)	-0.0949 (0.0875)	-0.113 (0.0886)
Alcohol	0.110 (0.168)	0.120 (0.172)	0.135 (0.172)	0.168 (0.179)	0.176* (0.0955)	0.185* (0.0966)	0.187* (0.0955)	0.193** (0.0966)
Age	-0.0113 (0.0135)	-0.0166 (0.0137)	-0.0124 (0.0134)	-0.0168 (0.0135)	-0.00210 (0.00755)	-0.00228 (0.00764)	-0.00179 (0.00748)	-0.00187 (0.00757)
Good access		0.467 (0.283)		0.427 (0.263)		0.273 (0.239)		0.238 (0.242)
Satisfied		-0.252 (0.277)		-0.286 (0.279)		0.126 (0.201)		0.147 (0.201)
Asset poor (bottom 2 quintiles)			0.114 (0.178)	0.156 (0.182)			0.0954 (0.0795)	0.0789 (0.0817)
Constant	0.974** (0.437)	1.047** (0.435)	0.887** (0.436)	0.997** (0.437)	0.293 (0.230)	0.311 (0.232)	0.281 (0.221)	0.292 (0.223)
Observations	57	57	57	57	153	151	153	151
R-squared	0.130	0.183	0.136	0.193	0.071	0.088	0.078	0.093

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 5.A3: Regression results for late access (narrow definition) by first pregnancy vs. second or later pregnancy

Variables	First pregnancy				Second or later pregnancy			
	(1) late	(2) late	(3) late	(4) late	(5) late	(6) late	(7) late	(8) late
Black	-0.0385 (0.0994)	-0.0427 (0.102)	-0.0290 (0.0997)	-0.0287 (0.102)	-0.0425 (0.0881)	-0.0425 (0.0881)	-0.0660 (0.0902)	-0.0660 (0.0902)
Grant	-0.255 (0.166)	-0.250 (0.169)	-0.237 (0.166)	-0.235 (0.170)	0.0212 (0.0928)	0.0212 (0.0928)	0.00817 (0.0921)	0.00817 (0.0921)
High school	-0.0985 (0.0963)	-0.109 (0.103)	-0.116 (0.0935)	-0.131 (0.0991)	-0.148 (0.0957)	-0.148 (0.0957)	-0.110 (0.0923)	-0.110 (0.0923)
Employed	0.102 (0.0974)	0.0825 (0.101)	0.0957 (0.0977)	0.0792 (0.101)	-0.0181 (0.0883)	-0.0181 (0.0883)	-0.00296 (0.0884)	-0.00296 (0.0884)
Asset poor	0.101 (0.101)	0.107 (0.104)			-0.0286 (0.0985)	-0.0286 (0.0985)		
Partner	-0.211** (0.103)	-0.211* (0.106)	-0.209** (0.103)	-0.211* (0.106)	0.0584 (0.0889)	0.0584 (0.0889)	0.0569 (0.0883)	0.0569 (0.0883)
Unwanted preg	0.0128 (0.106)	0.00527 (0.108)	0.0293 (0.105)	0.0224 (0.107)	0.0884 (0.0969)	0.0884 (0.0969)	0.0703 (0.0953)	0.0703 (0.0953)
Alcohol	0.191 (0.144)	0.201 (0.147)	0.205 (0.148)	0.210 (0.151)	0.0987 (0.0999)	0.0987 (0.0999)	0.112 (0.100)	0.112 (0.100)
Age	-0.00656 (0.0103)	-0.00691 (0.0104)	-0.00695 (0.0103)	-0.00716 (0.0105)	-0.00365 (0.00846)	-0.00365 (0.00846)	-0.00245 (0.00838)	-0.00245 (0.00838)
Good access		0.207 (0.265)		0.173 (0.263)	0.259 (0.208)	0.259 (0.208)	0.234 (0.209)	0.234 (0.209)
Satisfied		0.00246 (0.196)		0.0290 (0.194)	0.134 (0.244)	0.134 (0.244)	0.138 (0.243)	0.138 (0.243)
Asset poor (bottom 2 quintiles)			0.0806 (0.111)	0.0766 (0.113)			0.0959 (0.0939)	0.0959 (0.0939)
Constant	0.375 (0.237)	0.387 (0.240)	0.418* (0.233)	0.431* (0.237)	0.428 (0.273)	0.428 (0.273)	0.326 (0.261)	0.326 (0.261)
Observations	73	73	73	73	135	135	135	135
R-squared	0.177	0.187	0.170	0.179	0.074	0.074	0.081	0.081

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 5.A4: Regression results for late access (broad definition definition) by unwanted vs. wanted pregnancies

Variables	Unwanted pregnancies				Welcome pregnancies			
	(1) late2	(2) late2	(3) late2	(4) late2	(5) late2	(6) late2	(7) late2	(8) late2
Black	0.00710 (0.104)	0.0146 (0.111)	-0.00397 (0.113)	0.00254 (0.118)	-0.0327 (0.0793)	-0.0304 (0.0808)	-0.0277 (0.0802)	-0.0273 (0.0818)
Grant	0.0671 (0.112)	0.0612 (0.116)	0.0584 (0.105)	0.0492 (0.110)	-0.0399 (0.0958)	-0.0393 (0.0993)	-0.0408 (0.0967)	-0.0402 (0.100)
High school	-0.00521 (0.105)	-0.0159 (0.111)	0.0105 (0.106)	0.00257 (0.111)	-0.0457 (0.0818)	-0.0539 (0.0843)	-0.0605 (0.0797)	-0.0688 (0.0816)
Employed	-0.107 (0.102)	-0.0955 (0.117)	-0.108 (0.102)	-0.105 (0.114)	-0.0603 (0.0797)	-0.0630 (0.0809)	-0.0652 (0.0796)	-0.0684 (0.0807)
Asset poor	-0.0309 (0.130)	-0.0480 (0.143)			0.0669 (0.0812)	0.0714 (0.0832)		
Partner	0.0274 (0.0978)	0.0271 (0.100)	0.0366 (0.105)	0.0346 (0.107)	0.0936 (0.0826)	0.0833 (0.0845)	0.0947 (0.0835)	0.0824 (0.0857)
First preg	-0.0310 (0.124)	-0.0331 (0.127)	-0.0277 (0.133)	-0.0357 (0.138)	-0.0572 (0.0931)	-0.0636 (0.0949)	-0.0579 (0.0938)	-0.0638 (0.0955)
Alcohol	0.0664 (0.111)	0.0553 (0.117)	0.0683 (0.115)	0.0548 (0.123)	-0.134 (0.102)	-0.128 (0.104)	-0.129 (0.102)	-0.122 (0.104)
Age	-0.00740 (0.00896)	-0.00668 (0.00938)	-0.00775 (0.00894)	-0.00729 (0.00929)	0.00143 (0.00805)	0.000807 (0.00820)	0.000754 (0.00801)	0.000219 (0.00817)
Good access		-0.0395 (0.194)		-0.0160 (0.181)		-0.0474 (0.257)		-0.0571 (0.261)
Satisfied		0.0848 (0.190)		0.0712 (0.192)		0.142 (0.216)		0.161 (0.217)
Asset poor (bottom 2 quintiles)			0.0139 (0.119)	0.00221 (0.125)			0.0329 (0.0852)	0.0448 (0.0881)
Constant	1.120*** (0.290)	1.112*** (0.297)	1.096*** (0.291)	1.094*** (0.301)	0.716*** (0.245)	0.734*** (0.249)	0.764*** (0.236)	0.780*** (0.241)
Observations	57	57	57	57	153	151	153	151
R-squared	0.070	0.075	0.070	0.072	0.047	0.049	0.044	0.046

Standard errors in parentheses

*** p<0.01, ** p<0.05, *p<0.1

Table 5.A5: Regression results for late access (broad definition) by first pregnancy vs. second or later pregnancy

Variables	First pregnancies				Second or later pregnancies			
	(1) late2	(2) late2	(3) late2	(4) late2	(5) late2	(6) late2	(7) late2	(8) late2
Black	0.0531 (0.111)	0.0708 (0.114)	0.0566 (0.111)	0.0717 (0.113)	-0.0571 (0.0774)	-0.0571 (0.0774)	-0.0590 (0.0795)	-0.0590 (0.0795)
Grant	-0.374** (0.185)	-0.398** (0.189)	-0.361* (0.185)	-0.391** (0.189)	0.0568 (0.0815)	0.0568 (0.0815)	0.0576 (0.0812)	0.0576 (0.0812)
High school	0.0424 (0.108)	0.00635 (0.115)	0.0317 (0.104)	0.000190 (0.110)	-0.0591 (0.0840)	-0.0591 (0.0840)	-0.0614 (0.0813)	-0.0614 (0.0813)
Employed	-0.303*** (0.109)	-0.297** (0.113)	0.307*** (0.109)	-0.299** (0.113)	0.0237 (0.0775)	0.0237 (0.0775)	0.0240 (0.0779)	0.0240 (0.0779)
Asset poor	0.0678 (0.113)	0.0428 (0.117)			0.0158 (0.0865)	0.0158 (0.0865)		
Partner	-0.0648 (0.115)	-0.0874 (0.118)	-0.0638 (0.115)	-0.0878 (0.118)	0.154* (0.0780)	0.154* (0.0780)	0.155** (0.0778)	0.155** (0.0778)
Unwanted preg	0.190 (0.119)	0.185 (0.121)	0.200* (0.118)	0.192 (0.119)	0.157* (0.0851)	0.157* (0.0851)	0.159* (0.0840)	0.159* (0.0840)
Alcohol	-0.151 (0.161)	-0.173 (0.165)	-0.139 (0.165)	-0.165 (0.168)	-0.0502 (0.0877)	-0.0502 (0.0877)	-0.0486 (0.0884)	-0.0486 (0.0884)
Age	0.0110 (0.0115)	0.0116 (0.0116)	0.0107 (0.0115)	0.0114 (0.0116)	-0.00819 (0.00743)	-0.00819 (0.00743)	-0.00828 (0.00739)	-0.00828 (0.00739)
Good access		-0.111 (0.296)		-0.124 (0.293)	-0.0415 (0.183)	-0.0415 (0.183)	-0.0448 (0.184)	-0.0448 (0.184)
Satisfied		0.218 (0.219)		0.225 (0.216)	0.197 (0.215)	0.197 (0.215)	0.199 (0.214)	0.199 (0.214)
Asset poor (bottom 2 quintiles)			0.0631 (0.124)	0.0454 (0.126)			0.0121 (0.0828)	0.0121 (0.0828)
Constant	0.529* (0.266)	0.535* (0.268)	0.557** (0.261)	0.552** (0.263)	0.911*** (0.239)	0.911*** (0.239)	0.918*** (0.230)	0.918*** (0.230)
Observations	73	73	73	73	135	135	135	135
R-squared	0.214	0.227	0.213	0.227	0.081	0.081	0.081	0.081

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 5.A6: Regression results for timing of antenatal care access by pregnancy month

Variables	(1) months	(2) months	(3) months	(4) months
Black	-0.300 (0.247)	-0.349 (0.253)	-0.203 (0.245)	-0.255 (0.250)
Grant	-0.0796 (0.289)	-0.101 (0.295)	-0.0596 (0.293)	-0.0816 (0.299)
High school	-0.315 (0.245)	-0.339 (0.254)	-0.369 (0.253)	-0.400 (0.264)
Employed	-0.241 (0.250)	-0.231 (0.258)	-0.258 (0.252)	-0.258 (0.260)
Asset poor (bottom 2 quintiles)	0.538** (0.263)	0.474* (0.270)		
Partner	-0.113 (0.247)	-0.102 (0.254)	-0.133 (0.249)	-0.117 (0.256)
First preg	-0.0728 (0.300)	-0.127 (0.307)	-0.181 (0.297)	-0.232 (0.303)
Unwanted preg	0.596** (0.263)	0.549** (0.268)	0.623** (0.269)	0.576** (0.274)
Alcohol	0.375 (0.300)	0.306 (0.305)	0.312 (0.301)	0.247 (0.306)
Age	0.0166 (0.0239)	0.0161 (0.0244)	0.0177 (0.0242)	0.0166 (0.0247)
Good access		0.331 (0.589)		0.445 (0.591)
Satisfied		0.169 (0.554)		0.174 (0.561)
Asset poor			0.215 (0.262)	0.133 (0.270)
Constant	3.372*** (0.722)	3.497*** (0.737)	3.451*** (0.748)	3.616*** (0.762)
Observations	199	193	199	193
R-squared	0.116	0.112	0.099	0.098

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Chapter 6

Conclusion

In this dissertation I considered essays on the financing, user acceptability and delivery of healthcare in South Africa.

A well-functioning and accessible health system should be able to actively influence health outcomes and improve inequality in health outcomes, thereby also leading to a reduction in income inequality and poverty over the longer term. Individuals should have equality of opportunity to access and utilise the system, irrespective of their individual life circumstances. Health outcomes in various critical public health areas, however, allude to health system ineffectiveness in South Africa. The aim of the dissertation was therefore, through essays on the financing, user acceptability and delivery of healthcare, to examine sources of health system ineffectiveness and, in particular, explore how user acceptability of healthcare services influences system effectiveness.

This exploration required consideration of both the demand- and supply-sides. South Africa is a country characterised by high poverty and income inequality levels. This inequality also manifests in an extremely polarised healthcare market, with strong divisions between the public and private health sectors. The well-funded private health sector serves slightly less than a third of South Africans⁸¹, with the remainder of South Africans having to access healthcare services in the public sector.

On the demand-side, socio-economic vulnerability and social norms may inhibit health visits, partly due to the polarised nature of the healthcare market, thereby severely limiting the ability of the health system to influence health outcomes. The absence of a payment mechanism in the public sector, although increasing access to healthcare, leaves individuals with little leverage over the quality and performance of healthcare providers. Coupled with limited system transparency, and information on system performance, as well as the generally rare event of bypassing of healthcare facilities, individuals locked into only the public healthcare market may have little choice and control over the type and quality of services they have to utilise. The only remaining and very limited choice may therefore be whether or not to go to the clinic.

On the supply-side, health system effectiveness is determined by the interaction between health system users and healthcare providers. The broader medical and public health literature provide

⁸¹ Only 17% of South Africans have private health insurance but up to a third utilise private healthcare providers.

evidence of users being turned away at healthcare facilities, medication not being available and basic clinical protocols not being adhered to. These are strong signals of potential ineffectiveness.

The first essay considered the above themes in terms of private health insurance (the medical schemes market) and whether lower cost middle-market private health insurance options can help overcome the inequity between South Africa's highly polarised public and private healthcare markets. This is an important topic, not only for the private insurance sector, but also potentially for the public sector because the extension of health insurance to more South Africans can assist in alleviating the burden from the constrained public health sector.

In the second essay, I estimated the impact of the extension of private health insurance to more South Africans on total healthcare utilisation and provider choice. This was analysed in the context of the launch and roll-out of GEMS and within the larger context of South Africa's polarised private and public healthcare markets. It is an important research question internationally with the emphasis on UHC, but also locally because it could provide an upper bound benchmark for the expected demand response if the NHI were to provide cover for GP visits as currently proposed.

The third essay examined gender patterns and missed opportunities in health seeking amongst a sample of adults with possible TB symptoms (coughed ≥ 2 weeks) in the Western Cape. The analysis considered whether health seeking action was taken and examined provider choices, whether TB sputum samples were asked for in primary healthcare facilities, as well as the correlates of these choices and behaviour. This is important because it speaks directly to protocol implementation, clinical monitoring and the ability of the health system to play a role in stopping the spread of a contagious disease such as TB.

Lastly, in the fourth essay, I explored the causes of late antenatal care access amongst a sample of women who gave birth in public healthcare facilities in metropolitan Cape Town in the Western Cape. Although South Africa fares comparatively well in terms of women attending antenatal care and giving birth in healthcare facilities, a large percentage of women still attend antenatal care late. Early antenatal care access can save lives by facilitating the early detection and treatment of HIV and other risky health conditions which can potentially impact both the health of the mother and infant. Early antenatal care access is of pivotal importance in minimising maternal morbidity and mortality in South Africa. The chapter considered the self-reported causes and socio-economic and health system correlates of late antenatal care access.

The conclusions of each of the chapters are summarised below.

6.1 Chapter 2

Against the backdrop of the proposed NHI scheme and the move to UHC, this chapter examined the role of medical schemes in widening access to healthcare choice in South Africa. The existence of medical schemes as healthcare funding mechanism allow consumers greater choice in healthcare services and, at the least, allows for the bypassing of rationing mechanisms in the public sector.

In the first section of the chapter, I considered compositional changes in the post-apartheid medical schemes market and demonstrated that the labour market essentially still sets the limits of the market. I then reviewed international literature on the determinants of private health insurance membership relative to affordability, health risk and healthcare acceptability.

Pooled versions of the LFS (2000-2007), QLFS (2008-2011) and GHS (2002-2011) were used to explore associations between medical scheme membership and various socio-economic and demographic factors using LPMs. The analysis affirmed the importance of labour market skills levels, public sector employment, education and income as correlates of medical scheme membership. After controlling for these variables, race remained a large predictor of medical scheme cover, evidence of the enduring inequity that remains entrenched in medical scheme membership.

In the absence of a number of regulatory changes in the market primarily aimed at increasing the affordability of medical schemes, the size of the formal skilled labour market will continue to set the limits of private voluntary health insurance in South Africa. Unless growth constraining regulatory barriers in the medical schemes market, such as the nature of PMBs, the absence of a REF and cost minimisation through improved gatekeeping or more structured health referral, area addressed, the above correlates will continue to have a large negating influence on medical scheme membership.

I estimated that if one could succeed in lowering medical scheme membership costs to R300 per member, the mid-point value in the R200-R400 range of the CMS's recent discussion around low-cost (healthcare) benefit options, medical schemes membership could be extended from 17% to include up to 27% of the South African population. This will assist in alleviating pressure from an already over-burdened public healthcare system and improve equity in healthcare access. Failure to actively expand this market will, over the medium to long term, undermine equity in healthcare access in South Africa and continue to contribute to unequal health outcomes between race groups.

6.2 Chapter 3

The chapter considered the causal impact of access to private health financing on total healthcare utilisation and healthcare provider choice. A reading of the current causal empirical literature suggests that health insurance only affects healthcare utilisation in countries where the healthcare system is highly polarised. I investigated this hypothesis in the context of South Africa, where apartheid era policies left the country with a highly segmented healthcare system. In order to estimate the causal effect of health insurance, I exploited the exogenous variation in medical scheme coverage induced by the implementation of GEMS.

The identification strategy used aspects of difference-in-difference and IV estimators to identify the causal effect of health insurance on health seeking behaviour. Two datasets were used to test, firstly, the effect of the initial implementation of this policy in 2007 and, secondly, the effect of the continued roll-out between 2008 and 2012. The first strategy utilised the health insurance and industry information in the LFSs (the biannual LFS from 2000 to 2007 and the QLFS since 2008) and the detailed information on health services in the GHS. This approach was required because the LFSs do not contain any information on health seeking behaviour and the GHSs under-capture medical scheme membership during the crucial first years of GEMS. These two datasets were then combined to produce a two-sample 2SLS estimate of the causal effect of interest. The first-stage estimate was done using the LFS/QLFS data, after which the first-stage estimates were then applied to the GHS data in order to estimate the second-stage equation with the instrumented insurance variable.

The second strategy used the three waves (2008, 2010, 2012) of NIDS, to estimate the first and the second stage of the instrumented variable regression. NIDS is a nationally representative panel survey covering about 7,000 households. Because the first wave occurs in 2008, the data does not allow comparison with the pre-GEMS period and I therefore use the gradual GEMS roll-out and consequent increase in the likelihood of being a member of a medical scheme amongst public sector workers over the four-year window to capture the exogenous impact of GEMS.

The results showed that health insurance has a large effect on healthcare utilisation amongst South Africans, and that this effect is likely to be mainly due to the increased usage of perceived higher quality healthcare. This result was found to be highly robust irrespective of the choice of dataset, sample period, specification and estimator.

The estimates from the two estimation strategies found similar results and provide evidence of a very large impact of insurance cover on the utilisation of healthcare services (71%-82%). Insurance also increased the likelihood of using private providers and, in particular, private doctors (66%).

While problems with the quality and user acceptability of public sector services are increasingly acknowledged, this analysis clearly indicated the link between the increased client choice, user acceptability (and associated quality) offered by health insurance and quantity of healthcare consumed. Providing healthcare services of greater acceptability levels to those without medical scheme coverage is likely to boost the per capita number of visits, which can yield significant public health benefits.

The analysis also offered useful inputs for the planning process supporting the pilot and eventual launch of NHI. The large responses in utilisation and provider choice seen following the introduction of GEMS suggest that providing access to private doctors at no cost to the user (as some of the most basic scheme options of GEMS did) is likely to cause a large increase in the demand for healthcare. The analysis also showed that most of this will be directed towards private doctors. Although the demand response of the “treated” sub-group may over-estimate the response of poor South Africans that face more constraints in terms of transport and associated costs of private healthcare visits, the upper bound estimates can help set parameters for scenario-based forecasts to ensure adequate healthcare workforce planning. Given the pressure on public health budgets and the shortage of doctors and medical specialists, these estimates also highlight the need for effective and fair rationing strategies and renewed emphasis on gatekeeping to accompany other NHI health reforms.

6.3 Chapter 4

In this chapter, I considered the health seeking behaviour of adults with possible TB symptoms (coughed ≥ 2 weeks) from a gender perspective in the Western Cape. The detailed, step-by-step analysis of health seeking behaviour and choices, starting from the identification of TB-symptomatic adults, allowed for the identification of missed opportunities in the context of TB detection and treatment.

Two datasets were used: firstly, a pooled version of the nationally representative GHS from 2002 to 2011 to provide a larger context to TB prevalence in South Africa and, secondly, a TB prevalence survey of 30,017 adults which was conducted in 2010 as part of the larger ZAMSTAR intervention in eight high TB prevalent communities of the Western Cape.

The chapter started by reviewing literature on gender and health and gender and TB, respectively, and then moved to provide more information on the study context and methods. Next, I analysed the TB prevalence survey by using bivariate analysis to consider gender patterns in the TB health seeking cascade. The socio-economic correlates of TB-symptomatic adults were also considered. The

results of the data analysis were discussed and interpreted by referring to how the results relate to other study findings. The chapter was concluded by briefly referring to the policy implications of the findings.

In the study, only one third of survey participants (30.2%) indicated they sought help for TB symptoms (coughed ≥ 2 weeks) and only one fourth of those who coughed ≥ 2 weeks (24.1%) reported these symptoms at primary healthcare facilities. It is of concern that some of the TB cases in the sample went undetected. This is not unexpected given that in South Africa TB is diagnosed mainly through passive case finding⁸². However, it is problematic that adults specifically presenting for possible TB symptoms at primary healthcare facilities and who met the well-defined testing protocol criterion for TB of a cough of ≥ 2 weeks were not asked for sputum samples. Almost a third of adults (31.3%) in the sample who presented at primary healthcare facilities with cough symptoms consistent with possible TB were not asked for a sputum sample.

Despite being significantly more likely to seek care for their cough ($P < 0.05$), women were significantly more likely not to be asked for a sputum sample when they sought care at a primary healthcare facility ($P < 0.05$). This exacerbates the already marginal position of socio-economically vulnerable women who carry a disproportionate HIV burden. I can only speculate why women who presented with TB symptoms were less likely than men to be offered sputum tests. Possibly healthcare workers do not view women as the typical TB patient. This possibility needs to be explored and, if so, healthcare workers require sensitisation training to break down existing stereotypes of the typical TB patient. The achievement of gender equity in access to and use of TB services, at the least, requires adherence to existing well-defined clinical protocols. The weak adherence to these protocols points towards acceptability and quality issues in the health system.

6.4 Chapter 5

To explore the issue of late antenatal care access, I used the methodological approach of Solarin and Black (2013) and conducted a cross-sectional survey at four public sector labour wards in metropolitan Cape Town (Western Cape) between October and November 2014. A total of 221 women were interviewed.

The survey captured self-reported timing of first antenatal care access. Using univariate, bivariate and multivariate analysis in the form of linear probability models, I explored late antenatal care

⁸² Passive case finding is where patients or health system users are expected to seek care for symptoms in the health system, allowing the health system to detect the presence of disease. The opposite of passive case is active case finding where the health system more actively tries to detect disease by going to health system users through, for example, community-based interventions.

access relative to socio-demographic and health characteristics, risky behaviour correlates and health system factors. In particular, I explored late antenatal care access relative to the three dimensions of healthcare access: availability, affordability and acceptability.

A high percentage (96.4%) of women attended antenatal care. More than a quarter (26.3%) of women attended late (≥ 20 weeks/5 months). Of those who attended late, 48.2% indicated late recognition of pregnancy as the major reason for delayed attendance. Of women who attended antenatal care, 14.6% reported not being seen by a clinician or screened at their first visit. The majority of women surveyed (77.7%) indicated that their pregnancies were unplanned.

Unlike Solarin and Black (2013), I did not find evidence of major supply-side constraints in the study setting. Antenatal care availability, affordability and acceptability were not systematically associated with late antenatal care access. Almost 15.0% of women who attended antenatal care reported not having had a booking visit at their first visit to the clinic. However, the practice of turning women away from antenatal care clinics to return at a later time for their booking visits is not in accordance with Department of Health policy and may lead to loss to follow up (disappearance from the system) or late access.

Late access seemed to be predominantly associated with demand-side factors of socio-economic vulnerability and risky behaviour. Women who accessed antenatal care late (≥ 20 weeks/5 months) in their pregnancies were more likely to fall in the bottom two quintiles of the asset index ($P < 0.05$), more likely to not have completed matric ($P < 0.05$) and more likely to report any alcohol consumption during pregnancy ($P < 0.05$). Lastly, late access was also significantly associated with pregnancy experience, with women who reported a second or later pregnancy more likely to access antenatal care late ($P < 0.05$).

Socio-economic vulnerability as captured by falling lower in the asset index distribution (bottom three quintiles, $P < 0.1$) was also identified as a correlate of a broader definition of late access (≥ 12 weeks/3 months). Unwanted pregnancies also emerged as a correlate of late attendance for this access definition ($P < 0.05$).

Health policy needs to be informed by evidence of both the supply and demand factors that influence late antenatal care attendance. In this analysis, late antenatal care access was reported to be mainly due to late recognition of pregnancy. This, together with high levels of unplanned and unwanted pregnancies, points towards issues related to effective access to contraception which can be viewed as an upstream supply-side factor. Better access to pregnancy tests may help women to recognise pregnancy earlier, while more information and incentives to enable early health seeking

may also improve the timeliness of antenatal care access. The public healthcare system should play a role in providing better access to contraception, pregnancy tests and information.

6.5 What does this mean for the South African health system?

What does the research contained in this dissertation collectively imply for the South African health system? I consider this question by, firstly, identifying high-level conclusions for the demand side and, secondly, presenting a discussion of the high-level conclusions for the supply side.

Obstacles to identifying demand-side barriers to health seeking:

While I was able to draw on various papers on healthcare access frameworks, such as the work by Penchansky and Thomas (1981) and more recent work by Birch, Thiede and McIntyre (2009), the work presented in this dissertation should indicate the absence of deep theoretical economic models to help explain the health seeking behaviour of individuals, particularly individuals in developing countries. Seminal papers on the theoretical determinants of the demand for healthcare and healthcare financing, such as those by Arrow (1963) and Grossman (1972), were written in the context of countries with very different levels of development and disease burdens than South Africa or those of many other developing countries. These papers also do not generally take into account the nature of health systems in developing countries, where interaction with the public health system is often associated with long waiting times, poor staff attitudes and events such as drug stock-outs. Individuals in developing countries face very different incentives sets in deciding whether or not to seek healthcare.

The work presented in this dissertation was constrained by the limited availability of large datasets that contain socio-economic variables, health or disease-specific variables and detailed data on both the actual and self-reported health seeking behaviour of individuals. The last publically available data from a Demographic and Health Survey (DHS) in South Africa is from 1998. This survey was conducted at a time before the HIV/AIDS crisis reaches its full impact and before TB had spread to the extent visible in the data presented in this dissertation. While a DHS was conducted in 2002 and 2003, the full dataset was never officially released. This leaves South African health researchers with limited options to understand the complex nature of the interaction between the social determinants of health, health seeking behaviour and health outcomes.

If the full impact of social exclusion and social norms on health seeking behaviour is to be identified and explored, the health system reform process will have to ensure that much more detailed data is collected on nationally representative household and individual levels.

Obstacles to identifying supply-side deficiencies:

Given the nature of a low capability trap as described by Andrews, Pritchett and Woolcock (2013) in Chapter 1, it is possible to ask whether the South African health system is stuck in such a capability trap and, if so, how to move out of this trap. This dissertation identified various instances of missed opportunities in the health system, including weak adherence to TB testing protocols in primary healthcare facilities, women being turned away at primary healthcare facilities for antenatal screening and even seemingly low or poor access to contraception. These examples indeed allude to the presence of a capability trap where form or appearance takes primacy over function.

In order to definitively answer the question on whether the South African public health system is stuck in a capability trap, much more detailed data on what is occurring in healthcare facilities and, more specifically, what is occurring in the interaction between healthcare staff and patients, is required. Data on user acceptability simply based on the perceptions or experiences of health system users may be biased if users of the system have only been exposed to a system which delivers healthcare of lower acceptability levels. If users' expectations of healthcare have been shaped by a system that tends to deliver care of relatively low acceptability levels, they will not be able to accurately assess what they are missing out on. As objectively as possible and preferably quantitative measurement of acceptability (or quality) of healthcare services is required across large numbers of healthcare providers to provide a nuanced, representative picture of the state of South African public healthcare. However, even this measurement approach will be difficult to get right. One approach could be to test the knowledge of healthcare providers on various protocols and diseases. However, knowledge does not always translate into actual best practice if the incentive structure which supports such a translation is deficient.

The standardised patient method (see for example, Das and Hammer, 2014) is a methodological approach which allows for the quantitative measurement of quality of healthcare. According to Das and Hammer (2014), this is viewed as the "gold standard" methodology in the measurement of quality of healthcare. Community members are trained as actors to present at healthcare providers with a given set of symptoms. The performance of the provider is measured relative to expected behaviour and interaction given the nature of the symptoms and the "patient's" health condition (Hammer and Das, 2014: 11):

"The quality of medical advice is assessed by the time spent with patients, by providers' adherence to case-specific checklists of recommended care, the likelihood of correct diagnosis, and the appropriateness of treatment."

An accurate evidence-base of what is happening in South African public healthcare facilities and, in particular, primary healthcare facilities will require the implementation of measurement approaches such as the standardised patient method that rely on well-defined clinic protocols and expected best practice. The typical (currently used) approach of patient satisfaction surveys may be subject to reporting biases and the low expectations of health system users and will not be able to reveal the full extent of health system ineffectiveness. It is important, however, to caution that even a standardised patient method is not without difficulties. Deciding what dimensions of expected clinical interaction to include in a standardised patient script or study protocol can raise various difficult medical and health system questions.

The recently established Office of Health Standards Compliance, which is intended to monitor compliance with the National Core Standards for Health Establishments (see Department of Health, 2011b), will not be able to obtain a sufficiently nuanced picture of healthcare quality in South Africa if it only focuses on the measurement of easily measurable features⁸³. If the full nature of the delivery of public healthcare in South Africa is to be fully understood, the Office will have to delve into the interaction between healthcare providers and patients. Ultimately, this is the last frontier of healthcare access.

Without a detailed, nuanced understanding of the interaction between health system users and the system and the gaps which arise, it will not be possible to embark on a process of problem-driven iterative adaption, the remedy to a low capability trap proposed by Andrews, Pritchett and Woolcock (2013) (see Chapter 1). Solutions to South Africa's seemingly low health system effectiveness can only be proposed once a thorough understanding of the problems has been established.

Summary:

In summary, the main contribution of this dissertation was to consider how access to health services in the broad sense and, more specifically, how the user acceptability of healthcare services influences not only health seeking behaviour, but also influences the ability of healthcare services to impact health outcomes. This is of critical importance in embarking on any major healthcare reform process such as South Africa is doing through the implementation of NHI. Much more detailed data and a deeper understanding of these factors will be required if South Africa is to improve health system effectiveness and eventually health outcomes.

⁸³ This includes the absence or presence certain physical facilities, equipment and adherence to health standards and processes at a high level.

References

- Abrahams, N., Jewkes, R. and Mvo, Z. 2001. Health care-seeking practices of pregnant women and the role of the midwife in Cape Town, South Africa. *Journal of Midwifery and Women's Health*. 46(4): 240-247. July/August.
- African National Congress (ANC). 2009. National Health Insurance Plan for South Africa, prepared by the Health and Education Subcommittee of the National Executive Committee of the African National Congress. Project funded by the Bill and Melinda Gates Foundation. 16 February.
- African National Congress (ANC) General Council. 2010. Additional discussion documents. December 2010. Available at: <http://www.anc.org.za/docs/discus/2010/aditionalo.pdf> (accessed 24 September 2015).
- Ahmed, S., Li, Q., Liu, L. and Tsui, A.O., 2012. Maternal deaths averted by contraceptive use: an analysis of 172 countries. *Lancet*. 380: 111–25.
- Akerlof, G. 1970. The Market for 'Lemons': Qualitative uncertainty and the market mechanism. *Quarterly Journal of Economics*. 84: 485-500.
- Albanesi, S. 2011. Maternal Health and Fertility: An International Perspective. World Development Report Background Paper. Washington, DC: World Bank.
- Alkema, L., Kantorova, V., Menozzi, C. and Biddlecom, A. 2013. National, regional, and global rates and trends in contraceptive prevalence and unmet need for family planning between 1990 and 2015: a systematic and comprehensive analysis. *Lancet*. 11 May. 381: 1642–52.
- Allanson, E.R. and Pattinson, R.C., 2015. Quality-of-care audits and perinatal mortality in South Africa. *Bulletin of the World Health Organization (WHO)*. 93: 424–428.
- Amnesty International. 2014. Struggle for maternal health: barriers to antenatal care in South Africa. October.
- Andrews, M., Pritchett, L. and Woolcock, M. 2013. Escaping capability traps through problem driven iterative adaptation (PDIA). *World Development*. 51: 234–244.
- Angrist, J.D. and Pische, J.S. 2008. *Mostly harmless econometrics: An empiricist's companion*. Princeton, N.J.: Princeton University Press.

- Arneson, R. 2015. Equality of Opportunity. Stanford Encyclopedia of Philosophy. In Zalta, E.N. (ed.) (Summer 2015 Edition). Available at: <http://plato.stanford.edu/archives/sum2015/entries/equality-opportunity/> (accessed 17 September 2015).
- Arrow, K. 1963. Uncertainty and the welfare economics of medical care. *American Economic Review*. 53(5): 941-973.
- Aschengrau A. and Seage, G.R. 2014. *Essentials of Epidemiology in Public Health*. Burlington, M.A.: Jones and Bartlett Learning.
- Austin, J. F., Dick J.M. and Zwarenstein, M. 2004. Gender disparity amongst TB suspects and new TB patients according to data recorded at the South Africa Institute of Medical Research Laboratory for the Western Cape Region of South Africa. *International Journal of Tuberculosis and Lung Disease*. 8: 435-439.
- Ayles, H., Muyoyeta, M., Du Toit, E., Schaap, A., Floyd, S., Simwanga, M., Shanaube, K., Chishinga, N., Bond, V., Dunbar, R., De Haas, P., James, A., Gey van Pittius, N.C., Claassens, M., Fielding, K., Fenty, J., Sismanidis, C., Hayes, R.J., Beyers, N., Godfrey-Faussett, P. and ZAMSTAR team. 2013. Effect of household and community interventions on the burden of tuberculosis in southern Africa: the ZAMSTAR community-randomised trial. *Lancet*. 382: 1183-1194.
- Baicker, K., Taubman, S., Allen, H., Bernstein, M., Gruber, J.H., Newhouse, J.P., Schneider, E.C., Wright, B.J., Zaslavsky, A.M. and Finkelstein, A.N. for the Oregon Health Study Group. 2013. The Oregon experiment – effects of Medicaid on clinical outcomes. *New England Journal of Medicine*. 368(18): 1713-1722.
- Bamford, L. 2013. Maternal, newborn and child health. In Adarath, A. and English, R. (Eds.). *South African Health Review 2012/13*. Durban: Health Systems Trust.
- Banta, D. 2003. What is the efficacy/effectiveness of antenatal care and the financial and organizational implications? Copenhagen, WHO Regional Office for Europe. December. Health Evidence Network Report. Accessed: http://www.euro.who.int/_data/assets/pdf_file/0007/74662/E82996.pdf (accessed 5 May 2015).
- Beegle, K. de Weerd, J. and Dercon, S. 2010. Orphanhood and human capital destruction: is there persistence into adulthood? *Demography*. 47(1): 163-180.

- Begum, V., de Colombani, P., Das Gupta, S., Salim, A.H., Hussain, H., Pietroni, M., Rahman, S., Pahan, D. and Borgdorff, M.W. 2001. Tuberculosis and patient gender in Bangladesh: sex differences in diagnosis and treatment outcome. *International Journal of Tuberculosis and Lung Disease*. 5: 604-610.
- Bernard, D. and Selden, TM. 2006. Workers who decline employment-related health insurance. *Medical Care*. 44(5): 112-118.
- Besley T., Hall J. and Preston I. 1999. The demand for private health insurance: do waiting lists matter? *Journal of Public Economics*. 72: 155-181.
- Blaauw, D. and Penn-Kekana, L. 2010. Maternal Health. In: Padarath, A., Fonn, S. (Eds.). *South African Health Review 2010*. Durban: Health Systems Trust. 3-28.
- Black, V., Brooke, S. and Chersich, M. 2009. Effect of Human Immunodeficiency Virus treatment on maternal mortality at a tertiary center in South Africa: a 5-year audit. *Obstetrics & Gynecology*. 114(2, Part I): 292-299. August.
- Bloom, S.S., Lippeveld, T. and Wypij, D. 1999. Does antenatal care make a difference to safe delivery? A study in Uttar Pradesh, India. *Health Policy and Planning*. 14(1): 38-48.
- Booyesen, F., Van der Berg, S., Burger, R., Von Maltitz, M. and Du Rand, G. 2008. Using an asset index to assess trends in poverty in seven sub-Saharan African countries. *World Development*. 36(6): 1113–1130.
- Botha, E., Den Boon, S., Verver, S., Dunbar, R., Lawrence, K.A., Bosman, M., Enarson, D., Toms, I. and Byers, N. 2008. Initial default from tuberculosis treatment: how often does it happen and what are the reasons? *International Journal of Tuberculosis and Lung Disease*. 12(7): 820–823
- Breen, R. A., Leonard, O., Perrin, F.M., Smith, C.J., Bhagani, S., Cropley, I. and Lipman, M.C. 2008. How good are systemic symptoms and blood inflammatory markers at detecting individuals with tuberculosis? *International Journal of Tuberculosis and Lung Disease*. 12(1): 44-49.
- Breyer, F., Bundorf, M.K. and Pauly, M.V. 2011. Chapter 11: Health care spending risk, health insurance, and payment to health plans. In Pauly, M.V., McGuire, T.G. and Barros, P.D. (Eds.). *Handbook of Health Economics: Volume II*. Amsterdam: Elsevier North-Holland.
- Brown, C.A., Sohani, S.B., Khan, K., Lilford, R. and Mukhwana, W. 2008. Antenatal care and perinatal outcomes in Kwale district, Kenya. *BMC Pregnancy and Health*. 8:2 doi:10.1186/1471-2393-8-2.

- Bundorf, M.K., Herring, B., Pauly, M.V. 2010. Health risk, income and employment-based health insurance. *Forum for Health Economics and Policy*. 13(2): Article 13.
- Burger, R., Bredenkamp, C., Grobler, C. and Van der Berg, S. 2012. Have public health spending and access in South Africa become more equitable since the end of apartheid? *Development Southern Africa*. 29(5): 681-703. December.
- Buvinić, M., Medici, A., Fernández, E. and Torres, A.C. 2006. Chapter 10: Gender differentials in health. In Jamison D.T., Breman, J.G., Measham, A.R., et al. (Eds.). *Disease control priorities in developing countries*. Washington D.C.: World Bank. 195-210.
- Calvert, C. and Ronsmans, C. 2013. HIV and the risk of direct obstetric complications: a systematic review and meta-analysis. *PLoS One*. 8(10). e74848.
- Card, D., Dobkin, C. and Maestas, N. 2008. Impact of nearly universal insurance coverage on health care utilization: evidence from Medicare. *American Economic Review*. 98:5: 2242–2258.
- Carrolli G., Villar J., Piaggio, G., Khan-Neelofur, D., Gulmezoglu, M., Mugford, M., Lumbiganon, P., Farnot, U., Bergsjø, P. and WHO Antenatal Care Trial Research Group. 2001. WHO systematic review of randomised controlled trials of routine antenatal care. *Lancet*. 357(9268): 1565-70. May 19.
- Case, A. and Ardington, C. 2006. The impact of parental death on school outcomes: longitudinal evidence from South Africa. *Demography*. 43(3): 401-420.
- Childs, B. and Erasmus, D. 2012. Review of the South African Market for Hospital Cash Plan Insurance. Report prepared for the FinMark Trust. September. Available at: http://cenfri.org/documents/Health%20Workshop/HospitalcashPlanInMrktRev_12.pdf (accessed 3 September 2014).
- Christian, C.S. and Crisp, N. 2012. Management in the South African public health sector: An x-inefficiency perspective. *Development Southern Africa*. 29(5): 725-737.
- Claassens, M. M., Jacobs, E., Cyster, E., Jennings, K., James, A., Dunbar, R., Enarson, D.A., Borgdorff, M.W. and Beyers, N. 2013. Tuberculosis cases missed in primary healthcare facilities: should we redefine case findings? *International Journal of Tuberculosis and Lung Disease*. 17: 608-614.
- Claassens, M.M., du Toit E., Dunbar, R., Lombard, C., Enarson, D.A., Beyers, N. and Borgdorff, M.W. 2014. Tuberculosis patients in primary care do not start treatment. What role do health system delays play? *International Journal of Tuberculosis and Lung Disease*. 17(5):603–607.

- Combrink, J. 2009. Is the original mandate of equitable access to affordable benefits being realised? The 2009 Annual GEMS Symposium. Adding value to the Public Service, 26 to 27 October 2009.
- Conway, K.S. and Deb, P. 2005. Is prenatal care ineffective? Or, is the 'devil' in the distribution. *Journal of Health Economics*. 24(3): 489-513.
- Conway, K.S. and Kutinova, A. 2006. Maternal health: does prenatal care make a difference? *Health Economics*. 15(5): 461-488.
- Coovadia, H., Jewkes, R., Barron, P., Sanders, D. and McIntyre, D. 2009. The health and health system of South Africa: historical roots of public health challenges. *Lancet*. 374: 817-834. September 5.
- Costa, J. and Garcia, J. 2003. Demand for private health insurance: how important is the quality gap? *Health Economics*. 12: 587-599.
- Council for Medical Schemes. 2001. Annual report 2000-2001. Pretoria: Council for Medical Schemes.
- Council for Medical Schemes. 2002. Annual report 2001-2002. Pretoria: Council for Medical Schemes.
- Council for Medical Schemes. 2003. Annual report 2002-2003. Pretoria: Council for Medical Schemes.
- Council for Medical Schemes. 2004. Annual report 2003-2004. Pretoria: Council for Medical Schemes.
- Council for Medical Schemes. 2005. Annual report 2004-2005. Pretoria: Council for Medical Schemes.
- Council for Medical Schemes. 2006. Annual report 2005-2006. Pretoria: Council for Medical Schemes.
- Council for Medical Schemes. 2007. Annual report 2006-2007. Pretoria: Council for Medical Schemes.
- Council for Medical Schemes. 2008. Annual report 2007-2008. Pretoria: Council for Medical Schemes.
- Council for Medical Schemes. 2009. Annual report 2008-2010. Pretoria: Council for Medical Schemes.

- Council for Medical Schemes. 2010. Annual report 2009-2010. Pretoria: Council for Medical Schemes.
- Council for Medical Schemes. 2011. Annual report 2010-2011. Pretoria: Council for Medical Schemes.
- Council for Medical Schemes. 2012. Annual report 2011-2012. Pretoria: Council for Medical Schemes.
- Council for Medical Schemes. 2014. Annual report 2013-2014. Pretoria: Council for Medical Schemes.
- Council for Medical Schemes. 2015. Circular 37 of 2015: Request for proposal – benefit design and pricing of a low cost benefit option. Pretoria: Council for Medical Schemes.
- Culyer, A.J. 2001. Some theory and its policy implications. *Journal of Medical Ethics*. 27: 275–283.
- Currie, J. 2011. Inequality at birth: some causes and consequences. Richard T. Ely Lecture. *American Economic Review: Papers and Proceedings*. 101(3): 1-22.
- Das, J. and Hammer, J. 2014. Quality of primary care in low-income countries: facts and economics. *Annual Review of Economics*. 6: 525-553.
- Das, J., Hammer, J. and Leonard, K. 2008. The quality of medical advice in low-income countries. Policy Research Working Paper Series 4501. Washington D.C.: World Bank.
- Day, C. and Gray, A. 2014. Health and related indicators. In: Padarath A, English R. (Eds.). *South African Health Review 2013/14*. Durban: Health Systems Trust. 201-346.
- De Wagt, A. and Connolly, M. 2005. Orphans and the impact of HIV/AIDS in sub-Saharan Africa. *Food, Nutrition and Agriculture*. 34: 1–9.
- Department of Health. 2007. Guidelines for maternity care in South Africa: Pretoria: Department of Health.
- Department of Health. 2009. National tuberculosis management guidelines. Pretoria: Department of Health. Available at:
http://familymedicine.ukzn.ac.za/Libraries/Guidelines_Protocols/TB_Guidelines_2009.sflb.ashx
(accessed 17 August 2015).

- Department of Health. 2010. Clinical guidelines: PMTCT (prevention of mother-to-child transmission). Pretoria: Department of Health.
- Department of Health. 2011a. National health insurance in South Africa: policy paper. Pretoria: Department of Health.
- Department of Health. 2011b. "Towards quality for patients": National core standards for health establishments in South Africa. Pretoria: Department of Health.
- Department of Health. 2013. The 2012 National Antenatal Sentinel HIV and Herpes Simplex type-2 prevalence Survey, South Africa. Pretoria: Department of Health.
- Department of Health. 2014. National consolidated guidelines for the prevention of mother-to-child transmission of HIV (PMTCT) and the management of HIV in children, adolescents and adults. 24 December. Pretoria: Department of Health. Available at: <http://www.sahivsoc.org/upload/documents/HIV%20guidelines%20Jan%202015.pdf> (accessed 6 May 2015).
- Development Bank of South Africa (DBSA). 2008. A Roadmap for the Reform of the South African Health System. Draft final report. November. Midrand: Development Bank of South Africa.
- District Health Information System (DHIS) database [dataset]. 2012. Pretoria: National Department of Health.
- Doiron, D. Jones, G. and Savage, E. 2008. Healthy, wealthy and insured? The role of self-assessed health in the demand for medical insurance. *Health Economics*. 17: 317-224.
- Dong, Y. 2013. How health insurance affects health care demand – a structural analysis of behavioural moral hazard and adverse selection. *Economic Inquiry*. 51(2): 1324–1344. April.
- Dorrington, R.E. and Bradshaw, D. 2011. Maternal mortality in South Africa: lessons from a case study in the use of deaths reported by households in censuses and surveys. *Journal of Population Research*. 28: 49-73.
- Downe, S., Finlayson, K., Walsh, D. and Lavender, T. 2009. 'Weighing up and balancing out': a meta-synthesis of barriers to antenatal care for marginalised women in high-income countries. *British Journal of Obstetrics and Gynaecology*. 116: 518-529.
- Dupas, P. 2011. Health behaviour in developing countries. *Annual Review of Economics*. 3: 425-449.

- Eastwood, S.V. and Hill, P.C. 2004. A gender-focused qualitative study of barriers to accessing tuberculosis treatment in The Gambia, West Africa. *International Journal of Tuberculosis and Lung Disease*. 8:70–75.
- Econex. 2010a. Updated GP and Specialist Numbers for SA. Health Reform Note 7.
- Econex. 2010b. Rationing as a response to supply side constraints. NHI Note 5. January.
- Econex. 2012. Medical scheme expenditure on private hospitals. Occasional note. August. Stellenbosch, South Africa: Econex.
- Economic Research Southern Africa (ERSA). 2014. Panel: Health financing and regulatory reforms. ERSA Research Brief. Available at: http://www.econrsa.org/system/files/publications/research_briefs/research_brief_14.pdf (accessed 6 September 2015).
- Eighty20. 2009. Demand-side analysis of medical scheme coverage and access in South Africa. Document prepared for the FinMark Trust.
- Engelbrecht, B. and Crisp, N. 2010. Improving the performance of the South African health system. In Padarath, A. and Fonn, S. (Eds.). *South African Health Review 2010*. Durban: Health Systems Trust. 195-204.
- European Collaborative Study. 2007. Time to Undetectable Viral Load after Highly Active Antiretroviral Therapy Initiation among HIV-Infected Pregnant Women. *Clinical Infectious Diseases*. 44: 1647–1656.
- Evans, D.B., Marten, R. and Etienne, C. 2012. Universal health coverage is a development issue. *Lancet*. 380: 864-5.
- Evans, D.K. and Miguel, E. 2007. Orphans and schooling in Africa: a longitudinal analysis. *Demography*. February. 44(1): 35-47.
- Facione, N.C. and Facione, P.A. 2006. The cognitive structuring of patient delay in breast cancer. *Social Science & Medicine*. 63: 3137–3149.
- Fang, H., Kean, M. and Silverman, D. 2008. Sources of advantageous selection: evidence from the Medigap market. *Journal of Political Economy*. 116 (2): 303-350.

- Feldstein, M. 1995. The economics of health and health care: what have we learned? What have I learned? *American Economic Review*. Papers and proceedings of the hundredth and seventh annual meeting of the American Economic Association Washington, DC, January 6-8. 85(2): 28-31.
- Figlio, D., Guryan, J., Karbownik, K. and Roth, J. 2014. The effects of poor neonatal health on children's cognitive development. *American Economic Review*. 104(12): 3921-3955.
- Filmer, D. and Pritchett, L. 2001. Estimating wealth effects without expenditure data – or tears: an application to educational enrollments in the state of India. *Demography*. February. 38(1): 115-132.
- Finkelstein, A. and McGarry, K. 2006. Multiple dimensions of private information: evidence from the long-term care insurance market. *American Economic Review*. 96(4): 938-958.
- Finkelstein, A., Taubman, S., Wright, B., Bernstein, M., Gruber, J., Newhouse, J.P., Allen, H., Baicker, K., and the Oregon Health Study Group. 2012. The Oregon health insurance experiment – evidence from the first year. *Quarterly Journal of Economics*. 127(3): 1057-1106.
- Finlayson, K. and Downe, S. 2013. Why do women not use antenatal care services in low- and middle-income countries? A meta-synthesis of qualitative studies. *PLoS Medicine*. 10(1). E1001373.
- Fish, T. and Ramjee, S. 2007. Unaffordable medical scheme contributions: a barrier to access to private health cover in South Africa. *South African Journal of Business Management*. 38(3): 29-37.
- Fleurbaey, M. & Schokkaert, E. 2012. Equity in health and health care. In Pauly, M.V., McGuire, T.G., & Barros, P.P. (Eds.). *Handbook of Health Economics Volume II*. 1004-1092. Amsterdam: Elsevier (North-Holland).
- Getahun, H. and Raviglione, M. 2010. Active case-finding for TB in the community: time to act. *Lancet*. 376: 1205–1206.
- Gissler, M. and Hemminki, E. 1994. Amount of antenatal care and infant outcome. *European Journal of Obstetrics and Reproductive Biology*. 56(1): 9-14.
- Goddard, M. and Smith, P. 2001. Equity of access to health care services: theory and evidence from the UK. *Social Science & Medicine*. 53: 1149–1162.

- Golub, J.E., Mohan, C. I., Comstock, C.W. and Chaisson, R.E. 2005. Active case finding of tuberculosis: historical perspective and future prospects. *International Journal of Tuberculosis and Lung Disease*. 9:1183–1203.
- Gordan, P. 2014. Budget speech by the Minister of Finance of South Africa. 26 February. Cape Town, South Africa.
- Gosoni, G.D., Ganapathy, S., Kemp, J., Auer, C., Somma, D., Karim, F., and Weiss, M.G. 2008. Gender and socio-cultural determinants of delay to diagnosis of TB in Bangladesh, India and Malawi. *International Journal of Tuberculosis and Lung Disease*. 12(7): 848-55.
- Govender, V., Chersich, M.F., Harris, B., Alaba O., Ataguba, J.E., Nxumalo, N. and Goudge, J. 2013. Moving towards universal coverage in South Africa? Lessons from a voluntary government insurance scheme. *Global Health Action*. 6: 19253.
- Government Employees Medical Scheme (GEMS). 2007. Annual report. Government Employees Medical Scheme: Pretoria.
- Government Employees Medical Scheme (GEMS). 2008. Annual report. Government Employees Medical Scheme: Pretoria.
- Government Employees Medical Scheme (GEMS). 2009. Annual report. Government Employees Medical Scheme: Pretoria.
- Government Employees Medical Scheme (GEMS). 2010. Annual report. Government Employees Medical Scheme: Pretoria.
- Government Employees Medical Scheme (GEMS). 2011a. Annual report. Government Employees Medical Scheme: Pretoria.
- Government Employees Medical Scheme (GEMS). 2011b. Fact sheet. Available at: <https://www.gems.gov.za/default.aspx?bGEww6k/+Xlof842EylcPipQw15isdOYWU+WVXxA1xg=> (accessed 23 September 2015).
- Government Employees Medical Scheme (GEMS). 2012a. GEMS and its performance. Presentation to the Portfolio Committee on Public Service and Administration. 16 May 2012.
- Government Employees Medical Scheme (GEMS). 2012b. Annual report. Government Employees Medical Scheme: Pretoria.

- Government Employees Medical Scheme (GEMS). 2013. Annual report. Government Employees Medical Scheme: Pretoria.
- Graham, W. and Hussein, J. 2003. Measuring and estimating maternal mortality in the era of HIV/AIDS. Workshop on HIV/AIDS and adult mortality in Developing Countries. Population Division, Department of Economic and Social Affairs, UN Secretariat. New York. http://www.un.org/esa/population/publications/adultmort/GRAHAM_Paper8.pdf (accessed 1 April 2015).
- Greenacre, M and Blasius J. (Eds.). 2006. *Multiple correspondence analysis and related methods*. Florida: CRC Press.
- Grobler, C. and Stuart, I., 2007. Health care provider choice. *South African Journal of Economics*. 75(2): 327-350.
- Grossman, M. 1972. On the concept of health capital and the demand for health. *Journal of Political Economy*. 80(2): 223-255.
- Gulliford, M., Figueroa-Munoz, J., Morgan, M., Hughes, D., Gibson, B., Beech, R. and Hudson, M. 2002. What does 'access to health care' mean? *Journal of Health Services Research and Policy*. 7(3): 186-188.
- Haddrill, R., Jones, G.L., Mitchell, C.A. and Anumba, D.O.C., 2014. Understanding delayed access to antenatal care: a qualitative study. *BMC Pregnancy and Childbirth*. 14: 207.
- Halse, P., Moeketsi, N., Mtombeni, S., Robb, G., Vilakazi, T., and Wen, Y.F. 2012. The role of competition policy in healthcare markets. Pretoria: Competition Commission of South Africa.
- Hammer, J. 2015. The case for universal health care is weak. Ajay Shah's blog. Available at: <http://ajayshahblog.blogspot.co.za/2015/07/the-case-for-universal-health-care-is.html> (accessed 15 September 2015).
- Havemann, R. and Van der Berg. S. 2003. The demand for healthcare in South Africa. *Journal for Studies in Economics and Econometrics*. 27(3): 1-27.
- Hawkes, S. and Buse, K. 2013. Gender and global health: evidence, policy, and inconvenient truths. *Lancet*. 381: 1783-1787.
- Hoa, N.P., Chuc, N.T. and Thorson, A. 2009. Knowledge attitudes and practices about tuberculosis and choice of communication channels in a rural community in Vietnam. *Health Policy*. 90: 8-1.

- Hoffman R.M., Black V., Technau K., van der Merwe K.J., Currier, J., Coovadia, A., Chersich, M. 2010. Effects of highly active antiretroviral therapy duration and regimen on risk for mother-to-child transmission of HIV in Johannesburg, South Africa. *Journal of Acquired Immune Deficiency Syndrome*. 2010. 54:35-41.
- Høj, Da Silva. D., Hedegard, K., Sandström, A. and Aaby, P. 2003. Maternal mortality: only 42 days? *British Journal of Obstetrics and Gynaecology*. November. 110: 995-1000.
- Honda, A., Ryan, M., Van Niekerk, R. and McIntyre, D. 2014. Improving the public health sector in South Africa: eliciting public preferences using a discrete choice experiment. *Health Policy and Planning (Advance Access)*. Doi: 10.1093/heapol/czu038.
- Hoque, M., Hoque, E. and Kader, S.B. 2008. Audit of antenatal care in a rural district of KZN, South Africa. *South African Family Practice*. 50(3): 66.
- Jack, W. and Lewis, M. 2009. Chapter 1: Health investments and economic growth: macroeconomic evidence and microeconomic foundations. In Spence, M. and Lewis, M. (eds.). *Health and growth: Commission on growth and development*. Washington D.C: World Bank. 1-40.
- Jayachandran, S. and Lleras-Muney, A. 2009. Life expectancy and human capital investments: evidence from maternal mortality declines. *Quarterly Journal of Economics*. 124(1): 349–397.
- Jeffery, B.S., Tsuari, M., Pistorius, L.R., Makin, J. and Pattinson, R.C. 2000. The impact of a pregnancy confirmation clinic on the commencement of antenatal care. *South African Medical Journal*. 90(2): 153-156.
- Johansson, E., Long, N.H., Diwan, V.K. and Winkvist, A. 2000. Gender and tuberculosis control: Perspectives on health seeking behaviour among men and women in Vietnam. *Health Policy*. 52: 33–51.
- Johar, M., Jones, G., Keane, M.P., Savage, E. and Stavrunova, O. 2013. The demand for private health insurance: do waiting lists matter? – Revisited. Nuffield College Working Paper 2013-W09.
- Karim, F., Akramul, I., Chowdhury, A.M.R, Johansson, E. and Diwan, V.K. 2007. Gender differences in delays in diagnosis and treatment of tuberculosis. *Health Policy and Planning*. 22:329–334.
- Karim, S.S.A., Churchyard, G.J., Karim, Q.A. and Lawn, S.D. 2009. HIV infection and tuberculosis in South Africa: an urgent need to escalate the public response. *Lancet*. 374(9694): 921-933.

- Khan, M.S, Dar, O., Sismandis, C., Shah, K. and Godfrey-Fausset, P. 2007. Improvement of tuberculosis case detection and reduction in discrepancies between men and women by simple sputum-submission instructions: a pragmatic randomised controlled trial. *Lancet*. 369: 1955-1960.
- Khan, T. 2015. Low-cost medical schemes delayed. *Business Day*. 16 October. Available at: <http://www.bdlive.co.za/business/healthcare/2015/10/16/low-cost-medical-schemes-delayed> (accessed 21 February 2016).
- Kill, A. 2012. What characterises the privately insured in universal health care systems? A review of the empirical evidence. *Health Policy*. 106: 60– 75.
- King, D. and Mossialos, E. 2005 The determinants of private medical insurance prevalence in England, 1997-2000. *Health Services Research*. 40(1): 195–212.
- King, G., Gakidou, E., Imai K. Lakin, J., Moore, R.T., Nall, C., Ravishankar, N., Vargas, M., Téllez-Rojo, M.M., Ávila, J.E.H., Ávila, M.H. and Llamas, H.H. 2009. Public policy for the poor? A randomised assessment of the Mexican universal health insurance programme. *Lancet*. 373(9673): 1447–1454.
- Kirigia, J.M., Sambo, L.G., Nganda, B., Mwabu, G.M., Chatora ,R., and Mwase, T. 2005. Determinants of health insurance ownership among South African women. *BMC Health Services Research*. 5:17.
- Kulier, R., de Onis M., Gülmezoglu A.M. and Villar, J. 1998. Nutritional interventions for the prevention of maternal morbidity. *International Journal of Gynecology and Obstetrics*. 1998 December. 63(3): 231-46.
- Laththrop, E., Jamieson, D.J. and Danel, I. 2014. HIV and maternal mortality. *International Journal of Gynecology and Obstetrics*. 127: 213-215.
- Leibbrandt, M., Woolard, I., Finn, A. and Argent, J. 2010. Trends in South African income distribution and poverty since the fall of Apartheid. OECD Social, Employment and Migration. Working Papers, No. 101. Paris: OECD Publishing.
- Levine, D., Polimeni, R., and Ramage, I. 2014. Insuring health or insuring wealth? An experimental evaluation of health insurance in rural Cambodia. IRLE Working Paper No. 109-14. Available at: <http://irle.berkeley.edu/workingpapers/109-14.pdf> (accessed 9 September 2014).
- Lienhardt, C. 2001. From exposure to disease: the role of environmental factors in susceptibility to and development of tuberculosis. *Epidemiological Review*. 23(2): 288-301.

- Lienhardt, C., Rowley, J., Manneh, K., Lahai, G., Needham, D., Milligan, P. and McAdam, K.P.W.J. 2001. Factors affecting time delay to treatment in a tuberculosis control programme in a sub-Saharan African country: the experience of The Gambia. *International Journal of Tuberculosis and Lung Disease*. 5(3): 233-239.
- Lleras-Muney, A. 2005. The relationship between education and adult mortality in the United States. *Review of Economic Studies*. 72(1): 189-221.
- Long, N.H., Diwan, V.K. and Winkvist, A. 2002. Difference in symptoms suggesting pulmonary tuberculosis among men and women. *Journal of Clinical Epidemiology*. 55: 115-120.
- Long, N.H., Johansson, E., Diwan, V. K. and Winkvist, A. 2001. Fear and social isolation as consequences of tuberculosis in Vietnam: a gender analysis. *Health Policy*. 58: 69-81.
- Long, N.H., Johansson, E., Lönnroth, K., Eriksson, B., Winkvist, A. and Diwan, V.K. 1999. Longer delays in tuberculosis diagnosis among women in Vietnam. *International Journal of Tuberculosis and Lung Disease*. 3:388–393.
- MacPhail, C., Pettifor, A.E., Pascoe, S. and Rees, H.V. 2007. Contraception use and pregnancy among 15–24 year old South African women: a nationally representative cross-sectional survey. *BMC Medicine*. 5:31
- Målvik, M., Hoa, D.T.P, and Thomsen, S. 2012. Causes and determinants of inequity in maternal and child health in Vietnam. *BMC Public Health*. 12:641.
- Matsotso, M.P. and Fryatt, R. 2013. National Health Insurance: The first 18 months. In Padarath, A. and English, R. (Eds.). *South African Health Review 2012/13*. Durban: Health Systems Trust. 21-36.
- Mayosi, B.M., Flisher, A.J., Lalloo, U.G., Sitas, F., Tollman, S.M. and Bradshaw, D. 2009. Health in South Africa 4: The burden of non-communicable disease in South Africa. *Lancet*. 374: 934-47. September.
- McIntyre, D. 2010. National Health Insurance: providing a vocabulary for public engagement. In: Fonn, S. and Padarath, A., (Eds.). *South African Health Review 2010*. Durban: Health Systems Trust. 145-156.
- McIntyre, D., Goudge, J., Harris, B., Nxumalo, N. and Nkosi, M. 2009. Prerequisites for national health insurance in South Africa: results of a national household survey. *South African Medical Journal*. 99: 725-729.

- McIntyre, D., Thiede, M. and Birch, S., 2009. Access as a policy-relevant concept in low- and middle-income countries. *Health Economics, Policy and Law*. 4(2): 179-193.
- McIntyre, D., Thiede, M., Nkosi, M., Mutyambizi, V., Castillo-Riquelme, M. and Gilson, L. 2007. A critical analysis of the current South African health system. SHIELD Work Package 1 Report. Health Economics Unit, University of Cape Town; Centre for Health Policy, University of Witwatersrand. April.
- McKinley, J, and McKinley, S.J. 1997. The questionable contribution of medical measures to the decline of mortality in the United States in the twentieth century. *Millbank Memorial Fund Quarterly/Health and Society*. 55 (3): 404–28.
- McLaren, Z., Ardington, C. and Leibbrandt, M. 2013. Distance as barrier to healthcare access in South Africa. A Southern Africa Labour and Development Research Unit (SALDRU) Working Paper Number 97. Cape Town: SALDRU, University of Cape Town.
- McLaren, Z., Ardington, C. and Leibbrandt, M. 2014. Distance decay and persistent health care disparities in South Africa. *BMC Health Services Research*. 14:541.
- McLeod, H. and Grobler, P. 2008. Risk equalization and risk selection in South Africa. Paper prepared for the RAN meeting in 2008. 30 January.
- McLeod, H. and Grobler, P. 2010. Risk equalisation and voluntary health insurance: The South Africa experience. *Health Policy*. 98: 27-38.
- McLeod, H. and Ramjee, S. 2007. Medical schemes. In Harrison, S. Bhana, R. and Ntuli, A. (Eds.). *South African Health Review 2007*. Durban: Health Systems Trust. 47-70.
- McLeod, H. 2012. Affordability of health insurance. National Health Insurance Policy Brief 9 for Innovative Medicines South Africa (IMSA).
- Meer, J. and Rosen, H.S. 2004. Insurance and the utilization of medical services. *Social Science & Medicine*. 58: 1623–1632
- Meintjes, G., Schoeman, H., Morroni, C., Wilson, D. and Maartens, G. 2008. Patient and provider delay in tuberculosis suspects from communities with a high HIV prevalence in South Africa: A cross-sectional study. *BMC Infectious Disease*. 8: 72.
- Merenstein D.J., Schneider, M.F., Cox, C., Schwartz, R., Weber, K., Robison, E., Gandhi, M., Richardson, J. and Plankey, M.W. 2008. Association between living with children and adherence

- to highly active antiretroviral therapy in the Women's Interagency HIV Study. *Pediatrics*. 121: e787–793.
- Merenstein, D., Schneider, M.F., Cox, C., Schwartz, R., Weber, K., Robison, E., Gandhi, M., Richardson, J. and Plankey, M.W. 2009. Association of Child Care Burden and Household Composition with Adherence to Highly Active Antiretroviral Therapy in the Women's Interagency HIV Study. *AIDS Patient Care and STDs*. 23(4): 289-296.
- Miller, C.R., Davis J.L., Katamba, A., Sserwanga, A., Kakeeto, S., Kizito, F. and Cattamanchi, A. 2013. Sex disparities in tuberculosis suspect evaluation: a cross-sectional analysis in rural Uganda. *International Journal of Tuberculosis and Lung Disease*. 17: 480-485.
- Mills, A., Ataguba, J.E., Akazili, J., Borghi, J. Garshong, B., Makawia, S., Mtei, G., Harris, B., Macha, J., Meheus, F. and McIntyre, D. 2012. Equity in financing and use of health care in Ghana, South Africa, and Tanzania: implications for paths to universal coverage. *Lancet*. 380 (9837): 126-133.
- Moloabi, S. 2013. GEMS: Working harder to make a difference for you. 21 June 2013 Available at: www.gems.co.za (accessed 30 October 2014)
- Moodley, J. (Ed.). 2014. Saving Mothers 2011-2013: Sixth report on confidential enquiries into maternal deaths in South Africa, Short report. Pretoria: Department of Health.
- Moran, N.F. and Moodley, J. 2012. The effect of HIV infection on maternal health and mortality. *International Journal of Gynecology and Obstetrics*. 119: S26-S29.
- Morrone, C. and Moodley, J. 2006. The role of urine pregnancy testing in facilitating access to antenatal care and abortion services in South Africa: a cross-sectional study. *BMC Pregnancy and Childbirth*. 6:26.
- Motsoaledi, A. 2014. Health budget vote speech by the Minister of Health. 23 July 2014. Cape Town, South Africa.
- Myer, L. and Harrison, A. 2003. Why do women seek antenatal care late? Perspectives from rural South Africa. *Journal of Midwifery and Women's Health*. 48: 268–272.
- Myer, L., Zullier, R., Bekker, L.G. and Abrahms, E. 2012. Systemic delays in the initiation of antiretroviral therapy during pregnancy do not improve outcomes of HIV-positive mothers: a cohort study. *BMC Pregnancy and Childbirth*. 12:94.

- National Department of Health, 2012. National Campaign on Accelerated Reduction in Maternal and Child Mortality (CARMMA) Dashboard (DHIS Indicators).
- Neyrolles, O. and Quintana-Murci, L. 2009. Sexual inequality in tuberculosis. *PLoS Medicine*. 6: e1000199.
- Ngomane, S. and Mulaudzi, F.M. 2012. Indigenous beliefs and practices that influence the delayed attendance of antenatal clinics by women in the Bohlabele district in Limpopo, South Africa. *Midwifery*. 28(1): 30-38.
- Nkonki, L., Chopra, M., Doherty, T., Jackson, D. and Robberstad, B. 2011. Explaining household socio-economic inequalities using multiple methods in three diverse settings in South Africa. *International Journal for Equity in Health*. 10: 13.
- Nyman, J.A. 1999. The economics of moral hazard revisited. *Journal of Health Economics*. 18: 811-824.
- O'Brien, R., Hunt, K., and Hart, G. 2005. 'It's cavemen stuff, but that is to a certain extent how guys still operate': men's accounts of masculinity and help seeking. *Social Science and Medicine*. 61(3): 503-516.
- O'Connell, T.S., Bedford, K.J., Thiede, M. and McIntyre, D. 2015. Synthesizing qualitative and quantitative evidence on non-financial access barriers: implications for assessment at the district level. *International Journal for Equity in Health*. 14:54.
- Palmer, N., Mills, A., Wadee, H., Gilson, L. and Schneider, H. 2002. A new face for private providers in developing countries: What implications for public health? Mimeo. London: London School of Hygiene and Tropical Medicine.
- Pattinson R.C. 2013 (Ed.). Saving Infants 2010 - 2011: Eighth Report on Perinatal Care in South Africa. Pretoria: Tshepesa Press, 2013. <http://www.ppip.co.za/wp-content/uploads/Saving-Infants-2010-2011.pdf> (accessed 6 May 2015).
- Pattinson, R. (Ed.). 2012. Saving Mothers 2008-2010: Fifth report on the confidential enquiries into maternal deaths in South Africa. Pretoria: Department of Health.
- Pattinson, R., Fawcus, S. and Moodley, J. (Eds.). 2013. Tenth interim report on confidential enquiries into maternal deaths in South Africa: 2011 and 2012. Pretoria: Department of Health.

- Pauly, M. 1968. The economics of moral hazard: comment. *American Economic Review*. 58(3): 531-537.
- Penchansky, R. and Thomas, J.W. 1981. The concept of access: definition and relationship to consumer satisfaction. *Medical Care*. 19(2): 127-140.
- Promoting equity through sex-specific medical research [Editorial]. 2014. *Lancet*. 383: 928.
- Propper, C. 1989. An econometric analysis of the demand for private health insurance in England and Wales. *Applied Economics*. 21: 777-92.
- Propper, C. 1993. Constrained choice sets in the U.K. demand for private medical insurance. *Journal of Public Economics*. 51: 287-307.
- Propper, C., Rees, H. and Green, K. 2001. The demand for private medical insurance in the UK: a cohort analysis. *Economic Journal*. 111: C180–C200.
- Quimbo, S., Florentino, Jhiedon, Peabody, J.W., Shimkhada, R., Pabelo, C. and Solon, O. 2008. Underutilization of social insurance among the poor: evidence from the Philippines. *PLoS One*. 3(10): e3379.
- Raatikainen, K., Heiskanen, N. and Heinonen, S. 2007. Under-attending free antenatal care is associated with adverse pregnancy outcomes. *BMC Public Health*. 7:268. doi:10.1186/1471-2458-7-268.
- Ramjee, S., Vieyra, T., Abraham, M., Kaplan, J. and Taylor, R. 2014. National Health Insurance and South Africa's private sector. In Padarath A, English R (Eds.). *South African Health Review 2013/2014*. Durban: Health Systems Trust. 93-104.
- Ramsay, A., Bonnet, M., Gagnidze, L., Githui, W., Varaine, F. and Guerin, P. J. 2009. Sputum sex and scanty smears: new case definition may reduce sex disparities in smear-positive tuberculosis. *International Journal of Tuberculosis and Lung Disease*. 13: 613-619.
- Ray, D. 1998. *Development Economics*. Princeton, NJ: Princeton University Press.
- Republic of South Africa, 1999. Regulations to the Medical Schemes Act No. 131 of 1998. GNR.1262 of 20 October 1999: Regulations. Pretoria: Republic of South Africa.
- Republic of South Africa. 2013. Millennium Development Goals country report 2013. Republic of South Africa: Pretoria.

- Republic of South Africa. 2015. National Health Insurance for South Africa: Towards Universal Health Coverage. Pretoria: National Department of Health.
- Schnippel, K., Mongwenyana, C., Long, L.C. and Larson, B.A. 2015. Delays, interruptions, and losses from prevention of mother-to-child transmission of HIV services during antenatal care in Johannesburg, South Africa: a cohort analysis. *BMC Infectious Diseases*. 15: 46. Published online 6 February.
- Schwartz, S.R., Rees, H., Mehta, S., Venter W.D.F., Taha, T.E. and Black, V., 2012. High incidence of unplanned pregnancy after antiretroviral therapy initiation: findings from a prospective cohort study in South Africa. *PLoS One*. 7(4): e36039.
- Sen, G., Östlin, P. and George, A. 2007 Unequal, unfair, ineffective and inefficient. Gender inequity in health: why it exists and how we can change it. Final report to the WHO Commission on Social Determinants of Health. September.
- Silal, S.P., Loveday P.K., Barnighausen, T. and Schneider, H. 2014. Local level inequalities in the use of hospital-based maternal delivery in rural South Africa. *Globalization and Health*. 10: 60.
- Söderlund, N. and Hansl, B. 2000. Health insurance in South Africa: an empirical analysis of trends in risk pooling and efficiency following deregulation. *Health Policy and Planning*. 15(4): 378-385.
- Solarin, I. and Black, V. 2013. "They Told Me to Come Back": women's antenatal care booking experience in inner-city Johannesburg. *Maternal and Child Health Journal*. 17:359-367.
- Sprague, C., Chersich, M.F. and Black, V. 2011. Health system weaknesses constrain access to PMTCT and maternal HIV services in South Africa: a qualitative enquiry. *AIDS Research and Therapy*. 8:10.
- Statistics South Africa. 2013. Mid-year population estimates 2013. Statistical Release P0302. Pretoria: Statistics South Africa.
- Statistics South Africa. 2014. Mortality and causes of death in South Africa, 2011: Findings from death notification. Statistical release P0309.3. 2014. Pretoria: Statistics South Africa.
- Stinson, K. and Myer, L. 2012. Barriers to initiating antiretroviral therapy during pregnancy: a qualitative study of women attending services in Cape Town, South Africa. *African Journal of AIDS Research*. 11(1): 65-73.

- Tapay, N. and Colombo, F. 2004. Private health insurance in OECD countries: the benefits and costs for individuals and health systems. *Towards high-performing health systems: Policy studies*. 265–319.
- Taylor, B., Taylor, A., Burns, D., Rust, J. and Grobler, P. 2007. Prescribed minimum benefits–quagmire or foundation for social health reform? *South African Medical Journal*. 97(6): 446–450.
- Thornton, R.L., Hatt, L.E., Field, E.M., Islam, M., Diaz, F.S., and González, M.A. 2010. Social security health insurance for the informal sector in Nicaragua: a randomized evaluation. *Health Economics*. 19(S1): 181-206.
- Thorson, A., Hoa, N.P. and Long, N.H. 2000. Health seeking behaviour of individuals with a cough of more than 3 weeks. *Lancet*. 356: 1823-3.
- United Nations, Department of Economic and Social Affairs, Population Division. 2014. *World Contraceptive Use 2014*. Geneva: United Nations.
- United Nations Children’s Fund (UNICEF). 2013. *Towards an AIDS-free generation: children and AIDS, sixth stocktaking report*. Geneva: United Nations Children’s Fund.
- Van den Heever, A. M. 2012. The role of insurance in the achievement of universal coverage within a developing country context: South Africa as a case study. *BMC Public Health*. 12(Suppl 1): S5.
- Van den Hof, S., Najlis, C.A., Bloss, E. and Stratemanns, M., 2010. A systematic review on the role of gender in tuberculosis control. Report prepared for Tuberculosis Control Programme (TB CAP). September.
- Van der Berg, S. and McLeod, H. 2009. Crude NHI plan threatens to make a bad situation worse. *Business Day*. 4 September. Available at: <http://www.bdlive.co.za/articles/2009/09/04/crude-nhi-plan-threatens-to-make-a-bad-situation-worse> (accessed 6 September 2015).
- Van Doorslaer, E., O’Donnell, O., Rannan-Eliya, R.P., Somanathan, A., Adhikari, S. R., Garg, C.C., Harbianto, D., Herrin, A.N., Huq, M.N., Ibragimova, S., Karan, A., CNg, C.W., Pande, B.R., Rachel Racelis, R., Tao, S., Tin, K., Tisayaticom, K., Trisnantoro, L., Vasavid, C. and Zhao, Y. 2006. Effect of payments for health care on poverty estimates in 11 countries in Asia: an analysis of household survey data. *Lancet*. 368: 1357–1364. October 14.
- Van Eeden, J. 2009. *The demand for healthcare and health insurance in South Africa*. Econex Research Note 8. February.

- De Vaus, D. 2002. *Surveys in social research*. London: Taylor and Francis.
- Vlassoff, C. 2007. Gender differences in determinants and consequences of health and illness. *Journal of Health, Population and Nutrition*. 25(1): 47-61.
- Von Fintel, D. 2007. Dealing with earnings bracket responses in household surveys - how sharp are midpoint imputations? *South African Journal of Economics*. 75 (2): 293-312.
- Wabiri, N., Chersich, M., Khangelani, Z., Blaauw, D., Goudge, J. and Dwane, N. 2013. Equity in maternal health in South Africa: analysis of health services access and health status in a National Household Survey. *PLoS One*. 8(9). E73864.
- Wagstaff, A. 2010. Estimating health insurance impacts under unobserved heterogeneity: the case of Vietnam's healthcare fund for the poor. *Health Economics*. 19: 189-208.
- Wagstaff, A., and E. van Doorslaer. 2003. Catastrophe and impoverishment in paying for health care: with applications to Vietnam 1993–98. *Health Economics*. 12: 921–34.
- Wang, J., Fei, Y., Hongbing, S. and Biao, X. 2008. Gender difference in knowledge of tuberculosis and associated health-care seeking behaviors: a cross-sectional study in a rural area of China. *BMC Public Health*. 8: 354.
- Watkins, R.E. and Plant, A.J. 2006. Does smoking explain sex differences in the global tuberculosis epidemic? *Epidemiology and Infection*. 134: 333-339.
- Weiss, M.G., Auer, C., Somma, D.B., Abouihia, A., Kemp, J., Jawahar, M.S., Karim, F., and Arias, N.L. 2006. Gender and tuberculosis: cross-site analysis and implications of a multi-country study in Bangladesh, India, Malawi, and Colombia. Geneva: WHO.
- Weiss, M.G., Somma, D., Karim, F., Abouihia, A., Auer, C., Kemp, J. and Jawahar, M.S. 2008. Cultural epidemiology of TB with reference to gender in Bangladesh, India and Malawi. *International Journal of Tuberculosis and Lung Disease*. 12: 837–847.
- Western Cape Government: Health. 2013. Annual performance plan 2013/2014. February. Western Cape Health Department: Cape Town.
- Western Cape Government: Health. 2014. Annual performance plan 2014/2015. February. Western Cape Health Department: Cape Town.
- Western Cape Government: Health. 2015a. Annual performance plan 2015/2016. February. Western Cape Health Department: Cape Town.

Western Cape Government: Health. 2015b. Maternity case records (booklet). Western Cape Health Department: Cape Town.

Western Cape Government: Health. 2015c. Clinic checklist: booking and follow-up of HIV negative patient. Western Cape Health Department: Cape Town.

Western Cape Government: Health. 2015d. Antenatal (pre-birth) health services. Available at: <https://www.westerncape.gov.za/service/antenatal-pre-birth-health-services> (accessed 4 May 2015).

Western Cape Government: Health: 2015e. Termination of pregnancy. Updated 9 March 2015. Available at: <https://www.westerncape.gov.za/service/termination-pregnancy-top> (accessed 30 September 2015).

Whitehead, M., Dahlgren, G. and Evans, T. Equity and health sector reforms: can low-income countries escape the medical poverty trap. *Lancet*. 358: 833-836. September.

World Bank. 2015. World Bank Development Indicators: Health Financing. Available at: <http://datatopics.worldbank.org/hnp/ThematicData> (accessed 23 September 2015).

World Health Organization (WHO). 2010. Health systems financing: the path to universal coverage. World Health Report 2010. Geneva: World Health Organization.

World Health Organization (WHO). 2013a. Global tuberculosis report 2013. Geneva, Switzerland: World Health Organization.

World Health Organization (WHO). 2013b. World Health Statistics. Geneva: World Health Organization.

World Health Organization (WHO). 2014. Global tuberculosis report 2014. Geneva: Switzerland: World Health Organization.

World Health Organization (WHO) Reproductive Health Library. 2014. Effect of partogram use on outcomes for women in spontaneous labour at term. 3 November. Available at: http://apps.who.int/rhl/pregnancy_childbirth/childbirth/routine_care/cd005461/en/ (accessed 29 September 2015).

World Health Organization, 2015a. Health expenditure ratios, by country, 1995-2013. Available at: <http://apps.who.int/gho/data/node.main.HEALTHEXPRATIOBYCOUNTRY?lang=en> (accessed on 23 September 2015).

World Health Organization (WHO). 2015b. Health statistics and information systems: maternal mortality (per 100,000 live births). Available at: <http://www.who.int/healthinfo/statistics/indmaternalmortality/en/> (accessed 20 April 2015).

World Health Organization (WHO). 2015c. World Health Statistics. Geneva: World Health Organization.

Xu, K., Evans, D.B., Kawabata, K., Zeramdini, R., Klavus, J., and Murray, C.J.L. 2003. Household catastrophic health expenditure: a multicountry analysis. *Lancet*. 362: 111–117. 12 July.

Young, R. 1996. The household context for women's health care decisions: impact of UK policy changes. *Social Science and Medicine*. 42(6): 949-963.

Datasets:

Southern Africa Labour and Development Research Unit. 2013a. National Income Dynamics Study 2008, Wave 1 [dataset]. Cape Town: Southern Africa Labour and Development Research Unit [producer]. Cape Town: DataFirst [distributor].

Southern Africa Labour and Development Research Unit. 2013b. National Income Dynamics Study 2010-2011, Wave 2 [dataset]. Cape Town: Southern Africa Labour and Development Research Unit [producer]. Cape Town: DataFirst [distributor].

Southern Africa Labour and Development Research Unit. 2013c. National Income Dynamics Study 2012, Wave 3 [dataset]. Cape Town: Southern Africa Labour and Development Research Unit [producer]. Cape Town: DataFirst [distributor].

Statistics South Africa. General Household Surveys 2002-2011 [datasets]. Pretoria: Statistics South Africa.

Statistics South Africa. Labour Force Surveys, September, 2000-2007 [datasets]. Pretoria: Statistics South Africa.

Statistics South Africa. Quarterly Labour Force Surveys, Quarter 3, 2008-2011 [datasets]. Pretoria: Statistics South Africa.

Appendix B to Chapter 5

Survey questionnaire

Women's experience of antenatal care in metropolitan communities in the Western Cape

SECTION 1: Demographic information

We are going to start with some general questions about you.

1. How old are you?

years old (*Only interview if 18 years or older*)

2. In which city, town, township do you usually live?

2a. Do you have another home (a place that you would call your home home)? Where is this?

3. How long have you lived in your current area?

- Less than 6 months
- More than 6 months but less than 1 year
- Between 1 year and 5 years
- More than 5 years
- My whole life

4. What is your country of birth?

5. How long have you lived in South Africa?

- My whole life
- Less than 6 months
- More than 6 months but less than 1 year
- Between 1 year and 5 years
- More than 5 years but not my whole life

6. How would you describe your marital status?

- Single
- Married civil/religious
- Married traditional/customary
- Polygamous marriage
- Living together as regular partners
- Never married
- Widow
- Separated
- Divorced

7. How would you describe yourself in terms of population group?

- Black
- Coloured
- Indian or Asian
- White
- Mixed
- Other
- I prefer not to say

8. What is your highest level of schooling completed?

- I did not go to school
- Primary School
- Up to grade 11
- Matric
- Diploma
- University degree or higher
- Other _____

9. How would you describe your main activity or work status best?

- Paid employee
- Paid family worker
- Self-employed
- Employer

- Unpaid family worker
- Schooling
- Unemployed

10. What are the income sources of your household? (more than one answer possible)

- Salary or wages
- Money from a business
- Investment income
- Pension income
- Remittances (money sent by people living elsewhere)
- Government grants, e.g. the old age grant, child grant, disability grant
- Sale of farming products and services
- Other sources of income, e.g. rental income
- No income

11. Which of the above income sources is your household's main source of income? (one answer only)

- Salary or wages
- Money from a business
- Investment income
- Pension income
- Remittances (money sent by people living elsewhere)
- Government grants, e.g. the old age grant, child grant, disability grant
- Sale of farming products and services
- Other sources of income, e.g. rental income
- No income

12. What is your approximate gross monthly or annual household income? Please include all sources. (one answer only)

- No income
- R1-R400
- R 401 - R 1,600
- R 1,601 - R 3,200
- R 3,201 - R 6,400
- R 6,401 - R8,000
- R8,001 or more
- I prefer not to say
- I don't know

13. Which of the following best describes where you live? (one answer only)

- House/townhouse
- Share a house/townhouse with other families
- Flat in block of flats
- House/flat/room in backyard (formal structure)
- Informal dwelling/ shack in backyard
- Informal dwelling/ shack NOT in backyard e.g. in informal/squatter settlement
- Room/ flatlet NOT in backyard but on a shared property
- Workers' hostel (bed/room)
- Other (specify)_____

14. Which of these best describes your housing situation? (one answer only)

- Own home (with or without a mortgage/loan)
- Rent
- Housing tied to job
- Occupied rent-free
- Other (specify)_____

15. How many rooms are there where you live (not including kitchen/toilet*)?

rooms (*Don't forget to check that ALL rooms are counted, not just bedrooms*)

16. How many people live in your house (including children*)?

people (*Don't forget to include newborn*)

17. Where do you obtain water for use in your home? (one answer only)

- Piped water inside the home
- Tap in the yard
- Water from a tap outside the yard
- Borehole
- Other (specify) _____

18. What type of toilet facility does your household use? (one answer only)

- Flush toilet connected to public sewerage system
- Flush toilet connected to a septic tank
- Chemical toilet
- Pit latrine with ventilation pipe
- Pit latrine without ventilation pipe
- Bucket toilet
- Other (specify)

19. Now I am going to ask you about the durable assets which you may or may not have in your house. Do you have at least one of the following in your house?

Fridge/freezer	Yes	No	Not sure
Hi-Fi stereo , CD player, MP3 player	Yes	No	Not sure
Radio	Yes	No	Not sure
Television	Yes	No	Not sure
Satellite dish	Yes	No	Not sure
Computer	Yes	No	Not sure
Cell phone	Yes	No	Not sure
Electric stove	Yes	No	Not sure
Paraffin stove	Yes	No	Not sure
Microwave	Yes	No	Not sure

Washing machine	Yes	No	Not sure
Motor vehicle in working condition	Yes	No	Not sure

SECTION 2: Pregnancy Experience

Now we are going to ask you some questions about your pregnancy and antenatal care.

20. How many times have you been pregnant? Please specify the exact number. (Include all pregnancies even those where the baby was not carried until birth for whatever reason.)

21. If this is not your first pregnancy, did you attend antenatal care for any of your other pregnancies? (exclude cases where healthcare was accessed for abortions)

- Yes
- No
- Not applicable

22. I am now going to ask you about recognising/knowing that you are pregnant. What do you think are the typical signs that you may be pregnant? (Instruction to interviewer: do not prompt the respondent by reading out answers, spontaneous mention only, more than one answer possible.)

- Missing/skipping a period/menstrual cycle
- Sore or swollen breasts
- Feeling very tired
- Feeling nauseous/morning sickness/vomiting
- Increased urination/going to the toilet more
- Craving certain foods
- Other _____

23. How did you confirm your current pregnancy? (one answer only, first wait for answer before reading options)

- I bought a pregnancy test
- I went to a clinic and was tested
- I just knew I was pregnant

- I missed a few periods/menstrual cycles
- I could feel the baby moving/kicking
- I went to the clinic for something else and found out I was pregnant
- Other _____

24. From whom did you seek care soon after you confirmed your pregnancy? (one answer only, first wait for answer before reading options)

- I went to see my GP
- I went to the antenatal clinic
- I went to other primary health facility
- I went to a traditional medicine provider
- I did not go to any health care provider
- Other _____

25. Did you go to antenatal care (at a government antenatal clinic) while you were pregnant with this baby?

- Yes...GO TO QUESTION 26
- No GO TO QUESTION 46

26. Did you attend antenatal care at more than one clinic?

- Yes...GO TO QUESTION 27
- No... GO TO QUESTION 28

27. What was your reason for attending more than one clinic?

28. About how many times did you go to the antenatal care clinic while you were pregnant with this baby? (Instruction to interviewer: please collect the exact number, provide the respondent enough time to think about the question.)

times

29. About how many times were you actually seen/checked by the nurse when you went to the clinic? (Instruction to interviewer: please collect the exact number, provide the respondent enough time to think about the question.)

times

30. If there were any times you went to the clinic and were not seen/checked by the nurse, what was the reason for this?

31. Which antenatal clinic did you go to (or was your main clinic if you attended more than one)?

- Clinic A
- Clinic B
- Clinic C
- Clinic D
- Clinic E
- Clinic D
- Clinic E
- Clinic F
- Clinic G
- Clinic H
- Clinic I
- Clinic J
- Clinic K
- Clinic L
- Clinic M
- Clinic N
- Clinic O
- Clinic P
- Clinic Q
- Other _____

32. How do you normally travel to the clinic? (one answer only)

- I walk there
- I use public transport, for example, a bus, train or minibus taxi
- I travel there in my own car
- Someone takes me in their car

Do not know

33. How long does it take to travel to the antenatal clinic when you use the mode of transport mentioned above (Specify for one direction only, using the usual/normal means of transport and include waiting times at the bus/taxi stop. One answer only.)

less than 15 minutes

15-29 minutes

30-89 minutes

90 minutes or more

Do not know

34. What is the typical cost of a one way trip to the clinic (ask only to respondents who indicated they use public transport in Q31)?

less than R10

R11-R20

R21-R30

R31-R50

More than R50

35. When did you FIRST go to the clinic (how far along in your pregnancy)? (one answer only)

(months)

Go to Q39

Less than 3 months (12 weeks) pregnant

More than 3 months (12 weeks) but less than 5 months (20 weeks) pregnant

Go to Q36

More than 5 months (20 weeks) but less than 7 months (28 weeks) pregnant

More than 7 months (28 weeks) but less than 8.5 months (34 weeks) pregnant

More than 8.5 months (34 weeks) pregnant

36. If you went to the clinic after you were approximately 4–5 months (about 20 weeks) pregnant, why did you not go earlier? (Tick all that apply – more than one answer possible)

I did not think it was necessary

I did not know I had to go any earlier than I did

I did not know where to go

I did not know I was pregnant until quite late

I did not have time to go any earlier

I kept putting it off

- I did not like the clinic
- I did go before 20 weeks but was turned away/told to come back at a later date
- Other _____

37. If you went to antenatal care after you were approximately 4–5 months (about 20 weeks) pregnant, are there any PERSONAL FACTORS that would have made you go sooner? (Tick all that apply – more than one answer possible)

- If I found out sooner that I was pregnant
- If I thought it was necessary
- If I was told or I knew I had to go sooner
- If I had time to go sooner
- If I thought there might be a problem with my baby
- If I did not feel well
- If I was encouraged to go by those around me
- If I had the support of those around me
- Nothing would have made me go sooner
- Other _____

38. If you went to antenatal care after you were approximately 4–5 months (about 20 weeks) pregnant, is there anything about the CLINIC that would have made you go sooner? (Tick all that apply – more than one answer possible)

- If the clinic staff were more friendly
- If I liked the clinic
- If I had money to go to the clinic
- If the clinic was closer to me
- If the clinic opening and closing times were better
- If I did not have to wait too long at the clinic
- If I did not have to be tested for HIV at the clinic
- I went earlier but I was turned away/told to come back
- Nothing about the clinic would have made me go sooner
- Other _____

39. If you did go for antenatal care, what were your main reasons for going? (Tick all that apply – more than one answer possible)

- I went because that's what you do when you are pregnant

- I went because you cannot get a bed in a labour ward without an antenatal clinic card
- I went because it is important for my own and the baby's health
- I went because I had a bad experience with one of my previous pregnancies and I thought going would help avoid this
- I went because I wanted to know my HIV status
- I went because I wanted to prevent HIV infection in my child
- Other _____

40. What happened the FIRST time you went to the clinic? (One answer only, then multiple answers possible for those who were seen by nurse)

Go to Q41

- I was asked to come back without being seen (i.e. checked by a nurse) or anything being done

Skip to Q45

- I was seen/checked by a nurse. The following was done:
 - I was asked about my medical/pregnancy history
 - My blood pressure was taken
 - My antenatal card/booklet was filled in
 - A blood test was done
 - I can't remember
 - None of the above
 - Other _____

41. If you were asked to come back another time without being seen (checked by a nurse), do you know why/were you told why?

- Yes _____ (GO TO NEXT QUESTION)
- No
- I can't remember

42. Why did you have to go back to the antenatal clinic at another time? (one answer only, only ask to those who say they were asked to come back)

- I did not have a booking appointment
- I did not have an identity document (ID)
- The clinic had stopped accepting new patients
- The clinic could not see any more patients on that day (quota system)
- The clinic only sees pregnant women in the mornings or afternoons or on specific days

- The clinic was closed
- The clinic workers were on strike/not at the clinic
- The clinic workers told me it was “too early” in my pregnancy
- Other reason (PLEASE CAPTURE REASON)

43. If you were asked to come back, how much later were you asked to come back?

Go to Q44

- Less than 2 weeks later
- More than 2 weeks but less than a month later
- More than a month later

44. If you were asked to come back to the clinic for your booking visit, how far along in your pregnancy at the time of your full booking visit?

months OR weeks

45. The FIRST time you were BOOKED at the clinic (i.e. seen/checked by a nurse) which of the following was done? (Tick all that apply – more than one answer possible. Ask this question to ALL WOMEN.)

- Some blood was taken for testing
- I was offered an HIV test
- I was offered proper HIV counselling before and after the HIV test
- I had a physical examination
- I was weighed
- My blood pressure was taken
- My antenatal card/booklet was filled in
- I was told what danger signs to be aware of during my pregnancy
- I was told how many times I would need to be seen at the clinic before I had my baby
- I was given information about how to be healthy while pregnant
- I was given information about nutrition
- I was given information about breast feeding
- Urine sample taken

Other _____

Ask questions 46 - 48 only to participants who did NOT attend antenatal care (at an antenatal care clinic) otherwise GO to question 48

46. Why did you not attend antenatal care? (Tick all that apply – more than one answer possible)

- I did not think it was necessary
- I did not know I had to go
- I did not know where to go
- I did not know I was pregnant until quite late
- I did not have time to go
- I kept putting it off till it was too late
- I did not like the clinic
- I was scared of being tested for HIV
- I had no one to look after my children
- I was turned away/told to come back
- Other (e.g. attended other healthcare provider) _____

47. If you did not attend antenatal care for this pregnancy, are there any PERSONAL FACTORS that would have made you go? (Tick all that apply – more than one answer possible)

- If I thought it was necessary
- If I was told or I knew I had to go
- If I had time to go
- If I thought there might be a problem with my baby
- If I did not feel well
- If I was encouraged to go by those around me
- If I had the support of those around me
- If I had someone to look after my children
- Nothing would have made me go
- Other (e.g. attended other healthcare provider) _____

48. If you did not attend antenatal care for this pregnancy, is there anything about the CLINIC that could be changed and that would make you decide to attend antenatal care in future? (Tick all that apply – more than one answer possible)

- If the clinic staff were more friendly
- If I liked the clinic
- If I had money to go to the clinic
- If the clinic was closer to me
- If the clinic opening and closing times were better
- If I did not have to wait too long at the clinic
- If the clinic did not test for HIV
- If the clinic provided proper counselling before and after the HIV test
- My previous experience with antenatal care was not very good
- I did go once but was turned away/told to come back
- Nothing about the clinic affected my decision
- Other _____

Now ask questions 49-58 to ALL RESPONDENTS irrespective of whether they attended antenatal care.

49. Was your current pregnancy planned?

- Yes, I wanted to get pregnant
- No, but I was happy when I found out I was pregnant
- No, and I was unhappy when I found out I was pregnant

The next few questions are about any complications or problems you may have had either while pregnant or during delivery. We also ask some questions about health choices.

50. Did you ever have to see a doctor while you were pregnant with this baby or with any of your previous pregnancies?

- Yes _____
(enter reason)
- No
- I don't know/I can't remember

51. Did you ever have any problems with this pregnancy? (for example, high blood pressure, bleeding, concerns about the growth of baby.)

- Yes _____

(enter problem)

- No
- I don't know/ I can't remember

52. Did you ever have any problems during a previous pregnancy? (for example, high blood pressure, bleeding, concerns about the growth of baby.)

- Yes _____

(enter problem)

- No
- I don't know/ I can't remember

53. Did you have any complications during this delivery?

- Yes _____

(enter response)

- No

54. Did you carry this pregnancy to term (to the end)?

- Yes, I was at least 37 weeks/8.5 months pregnant when I had this baby
- No, this baby was born early

55. How many weeks/months pregnant were you when this baby was born?

months OR weeks

- I don't know

56. Apart from antenatal care, did you seek care from any other health facility while you were pregnant?

- Yes...GO TO QUESTION 57
- No...SKIP QUESTION 57 AND GO TO QUESTION 58

57. Who or where did you seek care from? (more than one answer possible)

- General practitioner/doctor
- Primary health facility/private health clinic
- Pharmacist
- Traditional health provider
- Other _____

58. Have you ever had an HIV test?

- Yes
- No
- I don't know/I can't remember

59. During your pregnancy, how often did you drink alcoholic beverages (e.g. beer, wine, spirits, other)? (one answer only)

- Never
- Monthly or less
- 2 to 4 days a month (approximately once a week)
- 2 to 3 days a week
- 4 to 6 days a week
- 7 days a week

60. During your pregnancy, did you every drink alcoholic beverages (e.g. beer, wine, spirits, other) until you passed out?

- Yes
- No
- I don't know/I can't remember

END of interview for participants who did not attend ANC

SECTION 3: Experience of Antenatal Care

Instruction to interviewer – please ask these questions only to women who indicated they attended antenatal care in Q24

The following questions are about your experience of the care you received. Remember, your answers are **confidential**, please answer as honestly as you can.

Thinking about the care you received at the antenatal clinic, please state how much you agree or disagree with the following statements. (Please circle appropriate response)

61. I found the staff at the clinic very friendly.

Strongly agree Agree Undecided Disagree Strongly disagree

62. I thought the staff at the clinic behaved how health providers should behave.

Strongly agree Agree Undecided Disagree Strongly disagree

63. I thought the staff at the clinic knew what they were talking about.

Strongly agree Agree Undecided Disagree Strongly disagree

64. I thought the clinic looked clean.

Strongly agree Agree Undecided Disagree Strongly disagree

65. The amount of time I had to wait to be seen at the clinic was acceptable.

Strongly agree Agree Undecided Disagree Strongly disagree

66. I had no problem getting to the clinic.

Strongly agree Agree Undecided Disagree Strongly disagree

67. The staff provided me with all the information I needed.

Strongly agree Agree Undecided Disagree Strongly disagree

68. I thought the staff communicated very well with me.

Strongly agree Agree Undecided Disagree Strongly disagree

69. I was generally satisfied with the care I received.

Strongly agree Agree Undecided Disagree Strongly disagree

SECTION 4: Opinions and Suggestions

Instruction to interviewer – please ask these questions only to women who indicated they attended antenatal care in Q24.

70. Your experience of antenatal care is very important to us. We would like to know if you have any specific opinions about the care you received or if you have any suggestions about how to improve the care offered at the clinics.

END OF INTERVIEW

Instruction to interviewer: please record any interesting comments or experiences mentioned by the respondent here:
