AN ANALYSIS OF THE RELATIONSHIP BETWEEN HEALTH AND THE LABOUR MARKET IN SOUTH AFRICA

Chijioke Osinachi Nwosu

Student Number: NWSCHI002

Thesis presented for the degree of

DOCTOR OF PHILOSOPHY

in the

School of Economics

Faculty of Commerce

University of Cape Town

June, 2015

Supervisors

Prof. Ingrid Woolard

Assoc. Prof. Patrizio Piraino

This page is intentionally left blank

Plagiarism Declaration:

1. I know that plagiarism is a serious form of academic dishonesty.
2. I have read the document about avoiding plagiarism, am familiar with its contents and have avoided all forms of plagiarism mentioned there.
3. Where I have used the words of others, I have indicated this by the use of quotation marks.
4. I have referenced all quotations and properly acknowledged other ideas borrowed from others.
5. I have not and shall not allow others to plagiarise my work.
6. I declare that this is my own work.
7. I am attaching the summary of the Turnitin match overview (when required to do so).

Signature:

Signed by candidate

Declaration to be included in your thesis

In the front of your thesis, a signed and dated declaration in the following format must be included:

Declaration

I, Chijioke Osinachi Nwosu, hereby declare that the work on which this thesis is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university. I authorise the University to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever.

Signature:…… …. Date: …2 June, 2015….

ABSTRACT

The relationship between health and labour market outcomes is of academic and policy interest due to the essential role the labour market plays in engendering economic growth. It is in this regard that this thesis is both timely and essential especially in light of scant literature on the health-labour market relationship in South Africa. South Africa presents an interesting case for a study of this nature as it had experienced high disease burden and mortality, coupled with declining labour force participation in the period prior to this study. Furthermore, the relationship between health and labour market earnings as well as impairment-related wage discrimination is not well-known in South Africa. Therefore, this thesis sought to establish the relationship between health on the one hand, and labour force participation, wage determination and wage discrimination on the other, in South Africa. Data was obtained from the first and third waves of the National Income Dynamics Study (collected in 2008 and 2012 respectively), a rich and nationally representative survey dataset of South African households. Descriptive analysis and different econometric techniques like instrumental variables, censored quantile regression and Blinder-Oaxaca decomposition were used for estimation.

For the cross-sectional analysis, the study found significant impact of health on labour force participation of between 20% and 33% depending on the measure, while longer term relationships indicated statistically significant association (up to 11% for females and 16% for males). These figures indicate that the relationship between health and labour force participation was not just temporary. Males had higher labour force participation probability than females. Furthermore, grant receipt was associated with reduced labour force participation probability while education and age were associated with increased labour force participation. Also, marriage/cohabitation was negatively (positively) associated with female (male) labour force participation. In addition, labour force participation

probability was generally higher in other areas relative to traditional authority locations. These results conform to a priori expectations.

On the relationship (or gradient) between health and wages, the study established positive and statistically significant gradients between better physical, psychological and general health on the one hand, and wages on the other, among Africans and coloureds. This was even after controlling for education and other important wage determinants like occupational category, industry, union membership and gender. These gradients ranged from an elasticity of -0.06 to -0.07 for psychological health/depression to an elasticity of 0.31-0.45 for physical health (proxied by body mass index) in the short term. Also, persistently adverse general health and psychological conditions exhibited steep gradients.

Finally, the study found evidence of non-trivial impairment-related differences in returns to wage-determining characteristics (loosely termed wage discrimination) in both 2008 and 2012 for the average wage, while the proportion of estimated wage gaps contributed by impairment-related differences in returns increased over time. Similar findings were obtained across the wage distribution, as the proportion of total estimated wage gaps accounted for by returns to characteristics increased across waves in virtually all deciles of the wage distribution. Even in terms of magnitude, the returns/discrimination component of total estimated impairment-related wage gaps increased for most quantiles of the wage distribution. Finally, education and occupational class contributed the most to the explained wage gap across waves.

DEDICATION

To

My siblings

And the loving memory of my parents

ACKNOWLEDGEMENT

I thank God for his immense goodness in leading me thus far. Special thanks are due my supervisors, Prof. Ingrid Woolard and Assoc. Prof. Patrizio Piraino under whose guidance this thesis was written. Though difficult, the journey would have been unbearable without them. I am also grateful to Prof. Nicoli Nattrass who played an important role in the initial conceptualization of the thesis. I also thank colleagues at the School of Economics, University of Cape Town (UCT) for the insightful discussions and warm friendship we had throughout my stay there. Special mention is due for David Fadiran, Alfred Mukong, Sulemana Mahawiya and Grieve Chelwa. I am highly indebted to Mhbuba Shifa for painstakingly going through the thesis and offering useful suggestions. Dr. Opeoluwa Oyedele also deserves commendation for excellent research assistance.

In a special way, I am deeply grateful to the Carnegie Foundation; the Centre for Social Science Research, UCT; the African Economic Research Consortium; and the Postgraduate Funding Office, UCT for generous financial support throughout the Ph.D. journey. Your support contributed in no small way to my fulfilling research experience.

This acknowledgement will not be complete without a special mention for Prof. Hyacinth Ichoku and Dr. John Ataguba. You have been teachers, friends and role models whom I always look up to. You are always available whenever I call. I cannot thank you enough for the immense roles you have played in my life. Thank you.

To my siblings and parents, you have been the most important human enablers of my achievements so far. To my siblings, may God keep you safe and prosper your ways, and to my parents, rest secure in His bosom. Finally, to all who have brought me smiles, I express my heart-felt appreciation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ADL	Activities of daily living
ATE	Average treatment effect
BMI	Body mass index
CDF	Cumulative distribution function
CES-D8	Eight-question Center for Epidemiological Studies Depression Scale
CES-D10	Ten-question Center for Epidemiological Studies Depression Scale
CES-D20	Twenty-question Center for Epidemiological Studies Depression Scale
CSM	Continuing sample member
DALY	Disability-adjusted life year
GDP	Gross domestic product
HDI	Human development index
HIV/AIDS	Human immunodeficiency virus/Acquired immune deficiency syndrome
HRS	Health and retirement study
I	Impaired
IMR	Inverse Mill's ratio
IV	Instrumental variable
JMP	Juhn-Murphy-Pierce
LATE	Local average treatment effect
LFP	Labour force participation

LFS	Labour force survey
LPM	Linear probability model
NCD	Non-communicable disease
NEA	Not economically active
NI	Non-impaired
NIDS	National income dynamics study
OHS	October household survey
OLS	Ordinary least squares
PSID	Panel study of income dynamics
QLFS	Quarterly labour force survey
R	South African Rand
RCT	Randomized controlled trial
SA	South Africa
SAH	Self-assessed health
SALDRU	Southern Africa labour and development research unit
TB	Tuberculosis
TOT	Treatment effect on the treated
UCT	University of Cape Town (South Africa)
UK	The United Kingdom
UNDP	United Nations Development Programme
US	The United States of America
USD	The United States dollar

WHO	World Health Organization

CHAPTER 1

GENERAL INTRODUCTION

There is more agreement among economists regarding economic growth influencing health outcomes than the other way round (Mckeown, 1976; Pritchett & Summers, 1996). This forms the traditional view of the relationship between income and health among economists (Husain, 2010) and key channels through which income can affect health include affordable health care and increased food availability (Spence & Lewis, 2009). However, it has been increasingly recognized that health is not necessarily a mere end result of the growth process, but rather a key input into a country's growth function. It is in this sense that the World Health Organization's (WHO) Commission on Macroeconomics and Health states that extending the coverage of crucial health services to the poor could save many lives, reduce poverty, spur economic development and promote global security (WHO, 2001). This statement underlies the importance of health in generating and sustaining economic growth. More recently, it has been noted that mortality reductions have accounted for 11% of economic growth in low- and middle-income countries while the value of additional life years – a measure of the intrinsic value of better health – accounted for growth in full income of about 24% in low- and middle-income countries between 2000 and 2011 (Jamison et al., 2013). The foregoing therefore, shows that both income and health reinforce each other, an assertion supported by Deaton (2002).

The above assertion by the WHO notwithstanding, macro evidence of the economic impact of health appears muted at least in comparison to micro studies. The reasons for this weak relationship include weak institutions (which undermine the effect of health care spending), ineffective management and provider absenteeism (Spence & Lewis, 2009). On the other hand, Husain (2010) has noted that there seems to be more agreement on the issue of causality running from health to income from a microeconomic perspective (Savedoff & Schultz, 2000; Schultz, 2005;

Strauss & Thomas, 1998) compared to the more tenuous evidence from macro studies even though methodological and health measurement issues still plague many micro studies.

The labour market is an important institution which mediates the relationship between health and the economy. Currie and Madrian (1999) stressed the importance of examining the health-labour market relationship, noting that it is a means of evaluating the cost-effectiveness of interventions designed to cure or prevent diseases. Arguably, a more compelling reason to engage in such a study lies in the potential for good health to readily enhance an individual's ability to enjoy improved economic wellbeing. As Jack and Lewis (2009) observed, the most obvious reason why healthier people are more likely to be richer than their sicker counterparts is that they have a greater capacity to work harder, longer and more consistently than the latter. They enumerated a number of channels through which this relationship can come about. Firstly, healthy children are more likely to invest in human capital accumulation by completing more years of schooling than the relatively sick. Moreover, healthy children are less likely to experience illness-induced school absenteeism than the sick, thereby obtaining higher quality education than the latter. Such investments prepare them for future participation in the labour market and higher earnings. There is also an inter-generational dimension to this argument; most decisions affecting children are made by their parents on their behalf. Parental death through sickness is likely to diminish the probability of their children acquiring education and other requisite human capital in spite of social networks that exist in society (Case, Paxson, & Ableidinger, 2004; Jack & Lewis, 2009). Poorly educated kids then eventually earn lower incomes than the highly educated on the average when they enter the labour market.

Also, higher incomes from labour market participation, coupled with higher life expectancy, lead to higher savings which enhance physical capital accumulation necessary for growth. According to Bloom, Canning and Graham (2002), the Asian savings boom of the 1950-1990 period can be partly attributed to rising life expectancy. Also, Schultz (1979) observed that longer life spans incentivize people to invest more in education and firms in more on-the-job training, while the additional health capital and other forms of human capital result in higher productivity per worker. Moreover, the additional health capital also results in more years of labour force participation (LFP) and the reduction of time spent sick. Conversely, poor health leads to low productivity and reduction of time spent working among adults.

The foregoing has mainly considered the health-labour market relationship as largely operating through the supply of labour. However, it is possible that employers may fail to hire job seekers who have some impairment/disability, or pay workers with such characteristics less than their healthier counterparts even when such individuals have the same initial level of human capital as the non-disabled (Madden, 2004). Such lower participation/lower earnings of the impaired/disabled which operates through labour demand may be the result of health deterioration or due to discrimination. In this sense, discrimination occurs when otherwise productively identical individuals in similar jobs are treated differently in the labour market (in terms of employment, earnings, etc.) (Altonji & Blank, 1999). Thus, beyond ascertaining the relationship between health and earnings that is due to productivity differences, it is interesting from an equity and policy perspective to ascertain whether the earnings gap between the non-impaired and the impaired reflects discrimination against the latter.

South Africa presents an interesting case study for analysing the relationship between health and the labour market. Not only is evidence on this relationship largely lacking, the country has experienced high disease burden mainly due to the HIV/AIDS pandemic and other communicable diseases like tuberculosis (TB) as well as non-communicable diseases (NCDs) over the past decade. Indeed, the country has been described as suffering from a quadruple burden of disease due to mortality from communicable diseases, maternal and perinatal conditions and nutritional deficiencies; NCDs like diabetes and obesity; injuries; and HIV/AIDS[1] (Bradshaw et al., 2000). Furthermore, it has experienced declines in some key labour market indices like labour force participation at about the same time (Banerjee, Galiani, Levinsohn, McLaren, & Woolard, 2008; Statistics South Africa, 2012a). Moreover, the country is noted for institutionalized racial discrimination under the apartheid system. Understandably, this has resulted in a number of studies examining race-based wage gaps and discrimination (Chamberlain & Van der Berg, 2002; Mwabu & Schultz, 2000). However, studies on impairment/disability-based wage discrimination are virtually non-existent. Therefore, this thesis fills these research gaps by examining the relationship between health and the labour market in South Africa with particular focus on LFP, wages and unexplained wage gaps (loosely termed wage discrimination).

1.1 AIM AND OBJECTIVES

1.1.1 Aim

This thesis aims to ascertain the relationship between health and the labour market in South Africa.

1.1.2 Objectives

Specifically, I intend ascertaining:

A. Whether self-reported health (a proxy for actual health status) positively affects labour force participation in South Africa.
B. The magnitude of the gradient between health and wages over the wage distribution for physical, psychological and general health conditions.
C. The magnitude of impairment-related differences in returns to observable characteristics (loosely termed wage discrimination) in South Africa.

[1] Though HIV/AIDS conventionally belongs to the communicable disease group, its unusually high prevalence in South Africa led to its separate classification.

1.2 RATIONALE/JUSTIFICATION OF STUDY

This thesis is important in that it contributes to an important literature by examining how health relates to labour market outcomes. The importance of the labour market to economic growth cannot be over-emphasized and as a result, it is important to investigate factors necessary for improved labour market outcomes. As shown above, health can play an important role in fostering improved labour market outcomes, and ascertaining the nature of this relationship in South Africa will add to the richness of the debate on the determinants of labour market outcomes in the country.

1.3 DATA

Data for this thesis was obtained from the National Income Dynamics Study (NIDS). NIDS, a nationally representative panel dataset of South African households, is the result of an initiative by the South African Presidency to track changes in wellbeing of South Africans over time. It is a rich source of socio-economic and health indicators. The first wave was collected in 2008 while the second and third waves were collected in 2010 and 2012 respectively. Data collection for the fourth wave is currently underway. The analysis was conducted using the first and/or third wave(s) depending on the question answered. Wave 2 was excluded due to inconsistencies between it and other waves as well as Statistics South Africa data collected around the same period. Such inconsistencies include the inexplicable nontrivial drop in the unemployment rate as well as a dramatic increase in the number of individuals working less than ten hours weekly between waves 1 and 2. These inconsistencies have been documented earlier (Cichello, Leibbrandt, & Woolard, 2012). Data cleaning is ongoing and has resulted in some modification of publicly released data. Therefore, I used the most recent versions of the two waves as at the time of writing this thesis (i.e. version 5.1 of wave 1 and version 1.1 of wave 3).

This analysis was restricted to resident continuing sample members (CSMs), i.e. respondents who are resident in the household and are to be followed across waves. 7296 households with 28226 CSM resident household members were visited for interview in wave 1, while the respective numbers for wave 3 were 10236 and 30329. Of the 28226 wave 1 CSM household residents, 26776 were successfully interviewed while 21384 (79.9%) of these were successfully re-interviewed in wave 3. Given that 20% of those successfully interviewed in wave 1 were not re-interviewed in wave 3, wave 3 and dynamic analyses were generally weighted by panel weights to account for possible non-random attrition and complex survey design. Similarly, wave 1 analyses were suitably weighted with post-stratification weights to correct for complex survey design. The sample was further restricted to the working age population, defined here as individuals aged 20 to 56 years in wave 1. This age restriction was informed by the need to exclude many students and those likely to retire by wave 3 from the analysis especially as the statutory age of eligibility for old age pension is 60 years (the age restriction in Chapter 4 was 20 – 60 years in wave 3 given that most of the analysis in that chapter was based on wave 3 data). For instance, only 9% of those between 15 and 19 years in wave 1 were employed while 80% were economically inactive. Also, 81% of those above 60 years were economically inactive. The overall working age sample was therefore restricted to 12337 respondents made up of 6947 females and 5390 males. All monetary measures were deflated using Statistics South Africa's headline consumer price index (Statistics South Africa, 2015) with August 2012 as the base month. Throughout this thesis, I urge caution in interpreting any statistic involving the Asian/Indian population group (where presented) given the small number of respondents belonging to this racial category in the sample. A similar caveat has been issued elsewhere regarding this dataset (Ardington & Case, 2009). Also, though province has been shown elsewhere to be significantly associated with labour market outcomes like LFP in South Africa (Ntuli & Wittenberg, 2013; Winter, 1999), this thesis did not conduct province-based analysis given that the NIDS sample size is too small to guarantee accurate inference at provincial and district levels (National

Income Dynamics Study, 2013). For more detailed description of the data as well as survey methodology and weights, the interested reader should visit www.nids.uct.ac.za.

Finally, the exclusive use of NIDS is informed by the fact that other nationally representative South African datasets are either rich in labour market outcomes only (e.g. the Labour Force Surveys) or health outcomes only (e.g. the Demographic and Health Surveys). On the other hand, census data are impinged by the lack of detailed socio-economic indicators necessary in a study of this nature. Therefore, NIDS is unique for our purposes in that it is a rich source of labour market, health and other socio-economic indicators.

1.4 STRUCTURE OF THE REMAINING CHAPTERS

This thesis is divided into seven chapters. I give a brief overview of the South African economy and discuss the post-apartheid labour market in a general background to the study in Chapter 2. Also, a descriptive analysis of the labour market between 2008 and 2012 using the NIDS dataset is conducted. Chapter 3 provides a brief review of key theoretical models of labour supply, wage determination and wage discrimination, issues that are empirically investigated in subsequent chapters. Chapter 4 ascertains the impact of health on LFP. In this chapter, I use different models to try to recover policy-relevant causal estimates in an instrumental variables (IV) framework. But mindful of the challenges in obtaining purely exogenous instruments, I also obtain measures of association between health and LFP. Chapter 5 deals with examining the gradient between health and wages in South Africa. I make the case for such a study to focus on the relationship at different points of the wage distribution rather than concentrating on the mean only. This exercise is conducted using physical, psychological and general health indicators. In Chapter 6, I examine the existence of impairment-related wage discrimination in the South African labour market while I conclude the thesis in Chapter 7.

While the thesis has one theme (i.e. an analysis of the relationship between health and the labour market in South Africa), each of the three empirical chapters (i.e. chapters 4-6) is a "semi-autonomous unit" in that it has an introduction, empirical literature review, a "methods" section aimed at analysing each of the above specific objectives, results and discussion of such results and a conclusion. Finally, "significance" will denote statistical significance throughout this thesis. Also, "significance at conventional levels" denotes statistical significance with alpha set at 10% at most.

CHAPTER 2

2 BRIEF OVERVIEW OF THE POST-APARTHEID SOUTH AFRICAN ECONOMY AND LABOUR MARKET

2.1 INTRODUCTION

In this chapter, I provide a brief overview of the South African economy and a historical account of the post-apartheid labour market especially in relation to key issues analysed in this thesis. I show that group-related gaps in key labour market outcomes like LFP and wages are generally persistent over the post-apartheid period. Subsequently, I conduct a descriptive analysis of the NIDS dataset especially in relation to labour supply and wages as these are key issues to be examined in this thesis. This analysis is informed by the recognition that there are non-health factors which play an important role in determining labour market outcomes. Subsequently, I make the case that an incorporation of health into the analysis is likely to enrich labour market models. Therefore, this chapter examines the post-apartheid South African economy, analyses the historical and current state of the labour market and serves as an exploratory prelude toward a more holistic study of labour market outcomes in South Africa (i.e. the incorporation of health into the analysis) with particular focus on LFP, wage determination and wage discrimination.

2.2 BRIEF OVERVIEW OF THE SOUTH AFRICAN ECONOMY

South Africa emerged from the racially exclusive system of apartheid to a democratic society in 1994. Apartheid was characterized by the exclusion of other racial groups from equal participation in virtually every facet of the economy in preference to whites[2]. The population in 1996 was 40.5 million.

---

[2] There are four officially recognized racial groups in South Africa: Africans, coloureds, Asians/Indians and whites. Africans are indigenous black South Africans; coloureds are mainly of mixed origins; Asians/Indians are

This increased to 51.7 million by 2011 while the gender composition remained virtually unchanged over the period: the proportions of males and females in 1996 (2011) were 48.1% (48.7%) and 51.9% (51.4%) respectively (Statistics South Africa, 2012a). According to the 2011 population census, the racial make-up of the population was as follows: African- 79.2%, coloured- 8.9%, Asian- 2.5%, and white- 8.9% (Statistics South Africa, 2012a).

South Africa is an upper middle income country, one of few African countries to be so categorized. The economy recorded impressive growth before the 2008 global financial crisis. However, growth has slowed considerably afterwards. Real gross domestic product (GDP) growth rate increased by 2.3% between 1995 and 1999 while annual growth rates were 4.9%, 5.0%, 5.4% in 2004, 2005 and 2006 respectively. Growth however declined to 3.6% in 2011, 2.5% in 2012 and 1.9% in 2013, the latter, well below the sub-Saharan Africa average of 5.1% (African Economic Outlook, 2014; National Treasury, 2000; 2008; 2014). This growth decline has been attributed to sluggish growth in the country's European and North American trading partners as well as continued labour unrest especially in the manufacturing sector, among other factors (African Economic Outlook, 2014).

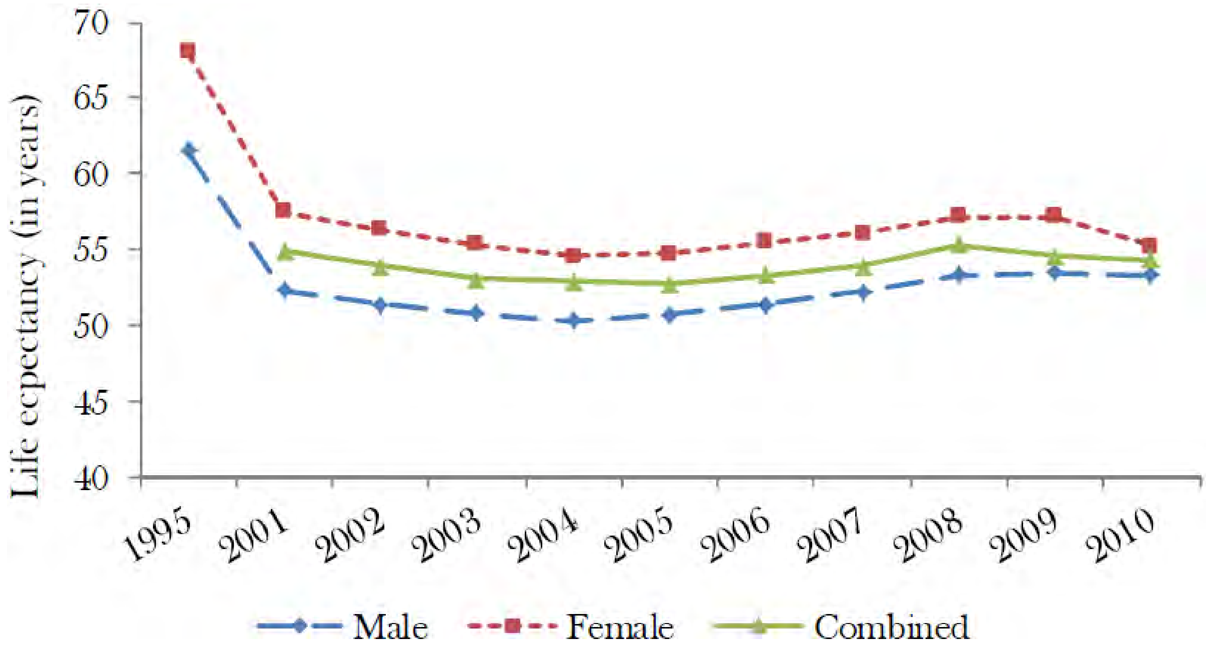
In terms of human development, South Africa ranked 118 out of 187 countries in the world in 2013 based on a Human Development Index (HDI) score of 0.658 (UNDP, 2014). This represents a decline from 0.724 in 1995, 0.695 in 2000; it was 0.619 in 2011 (du Toit, 2002; UNDP, 2011). This decline has been attributed to declining life expectancy attributed to the HIV/AIDS pandemic (Ataguba, 2012). Also, of the three components of the HDI: health, income and education, health contributed least to the index in 2011 while education had the highest share (UNDP, 2011).

South Africans of Asian (mainly Indian) origin; while whites are of European ancestry (Schultz & Mwabu, 1998b).

The country has been plagued with a high burden of disease especially in the first half of the last decade. As indicated in Chapter 1, South Africa has been described as suffering from an enormous burden of disease, while the HIV/AIDS pandemic and TB are responsible for a significant part of such high disease burden. The total number of AIDS deaths substantially increased from 198,030 in 2001 to a peak of 314,309 in 2006, while falling to 281,404 in 2010. Furthermore, per 100,000 of population, TB mortality increased from 133 in 1997 to 183 in 2000, and 230 in 2007 (UNICEF, 2010).

This high disease burden resulted in generally declining life expectancy between the 1990s and the past decade (see Figure 2.1 below). As noted by Ataguba (2012), average life expectancy fell from 65 years in 1995 to 54 years in 2010. Also, the apparent convergence of the gap in life expectancy between males and females can be attributed to disproportionately higher HIV/AIDS incidence among females.

Figure 2.1: Average life expectancy in South Africa, 2001-2010



Source: du Toit (2002); Development Indicators (2010), from Ataguba (2012)

South Africa is arguably one of the most unequal countries in terms of income, in the world. The country's Gini coefficient, an index of income inequality, has also increased over the post-apartheid period. Available survey data evidence shows that the Gini coefficient increased from 0.67 in 1993 to 0.70 in 2008, a 4.5% increase (Leibbrandt, Wegner, & Finn, 2011). This increase exhibited substantial variation across the different racial groups; Africans, coloureds, Asians and whites experienced Gini coefficient increments of 12.7%, 25.6%, 32.6% and 19.1% respectively. Such persistently high income inequality has been largely attributed to differences in labour market outcomes (Leibbrandt, Finn, & Woolard, 2012). Inter-racial differences are also clear when one considers household income differences along racial lines. Out of an estimated total household income of R1.57 trillion in 2010/2011, African households (who constituted more than three-quarters of the total number of households) earned only 44.6% of total annual household income while whites (who made up 12.4% of aggregate households) earned 40.1%. On the other hand, coloureds and Asians/Indians (8.5% and 2.5% of total households respectively) earned 9.9% and 5.4% of total household income respectively (Statistics South Africa, 2012c).

The foregoing has provided a snapshot of the South African economy especially with regard to key economic and health indicators. The next section focuses only on key labour market indicators like (un)employment, LFP, wages and wage gaps, issues that are very important in this thesis.

2.3 OVERVIEW OF LABOUR FORCE PARTICIPATION, EARNINGS AND WAGE DISCRIMINATION IN POST-APARTHEID SOUTH AFRICA

A cursory look at the post-apartheid South African labour market literature suggests the following, inter alia:

- Increasing female LFP, especially among Africans in the mid-90s to early 2000s (Casale & Posel, 2002) and declining aggregate LFP rates from around the mid-2000s (Banerjee et al., 2008; Statistics South Africa, 2012a)

- High levels of unemployment and poor labour absorption rates, especially among the African and coloured populations (Bhorat, 2004; Casale, 2004)
- Mainly involuntary and structural unemployment exacerbated by spatial rigidities with regard to accessing employment opportunities (Bhorat, 2004; Kingdon & Knight, 2004a)
- Skills-biased employment growth favouring the highly educated (Bhorat & Hodge, 1999; Bhorat, 2004)
- Substantial and sustained male wage premium (Wittenberg, 2014)
- Sustained race-based wage premiums in the following descending order: whites, Asians/Indians, coloureds and Africans (Wittenberg, 2014)
- Higher union premium among Africans relative to whites (Mwabu & Schultz, 2000; Schultz & Mwabu, 1998b)
- Wage discrimination driven by race, location and gender (Casale, 2004; Chamberlain & Van der Berg, 2002; Grün, 2004; Rospabé, 2002)

High rates of unemployment have remained a serious issue in post-apartheid South Africa especially as the unemployment rate increased over the period. This increase can be attributed to both labour supply and demand forces; partly due to higher LFP (relative to the apartheid era) as well as poor labour absorption rates, the latter being the result of higher population growth relative to job creation as well as skills mismatch.
With regard to demand-side factors, the South African economy has largely struggled, especially in the post-apartheid era, to create enough jobs for those entering the labour market. For instance, labour absorption rates fell from 56.3% in the 1990s to approximately 41% in 2010 (Ataguba, 2012; du Toit, 2002). In highlighting the deficit in the employment creation performance of the South African economy, Bhorat (2004) demonstrated that while the target employment growth rate – a measure of the desired rate

at which jobs are to be created in order to employ all new entrants into the labour market – grew by 52.4% between 1995 and 2002 (using the expanded definition), employment absorption was only 32% over this period. Even with the strict definition favoured by government agencies, employment absorption was still 40.4%. A breakdown of these figures by race and gender suggests wide variations along these characteristics. For instance, while the target employment growth rate among Africans was 66.9%, only 28% of these were absorbed into employment, while for whites, more than half of the 13% target employment growth rate got jobs over the period. For males, the target growth rate was 31.7% while 19.1% of these were absorbed, while for females, the target was 84.5% while 39.5% was absorbed.

Some of the factors that might have been responsible for such weak job creation outcomes include substantial wage growth especially for the unskilled and semi-skilled, skills-biased technological growth (partly as a consequence of phenomenal wage growth among the unskilled), as well as poor growth in the informal sector. Arndt and Lewis (2001), who reported that total employment (narrow definition) of the unskilled and semi-skilled workforce in 1999 was a mere 92% of their 1970 levels, largely attributed this to phenomenal growth in real remuneration per unskilled and semi-skilled worker relative to the highly-skilled. Indeed, 1999 semi-skilled and unskilled real wage levels grew to 250% of their 1970 levels while that of the highly skilled contracted to 90% of their 1970 levels. This perhaps significantly contributed to a substitution of capital for unskilled and semi-skilled labour over this period. A similar point was made by Bhorat and Hodge (1999) who noted that the rise in nominal wages in the mining sector in the 1980s shifted relative factor prices in favour of capital, thus increasing the capital intensity of extraction (indeed, they showed that capital-labour ratios substantially increased for all but one sector between 1970 and 1995). Given that capital accumulation often adversely affects the employment of unskilled labour relative to skilled labour, the unemployment rate among unskilled labour was exacerbated between 1970 and 1999.

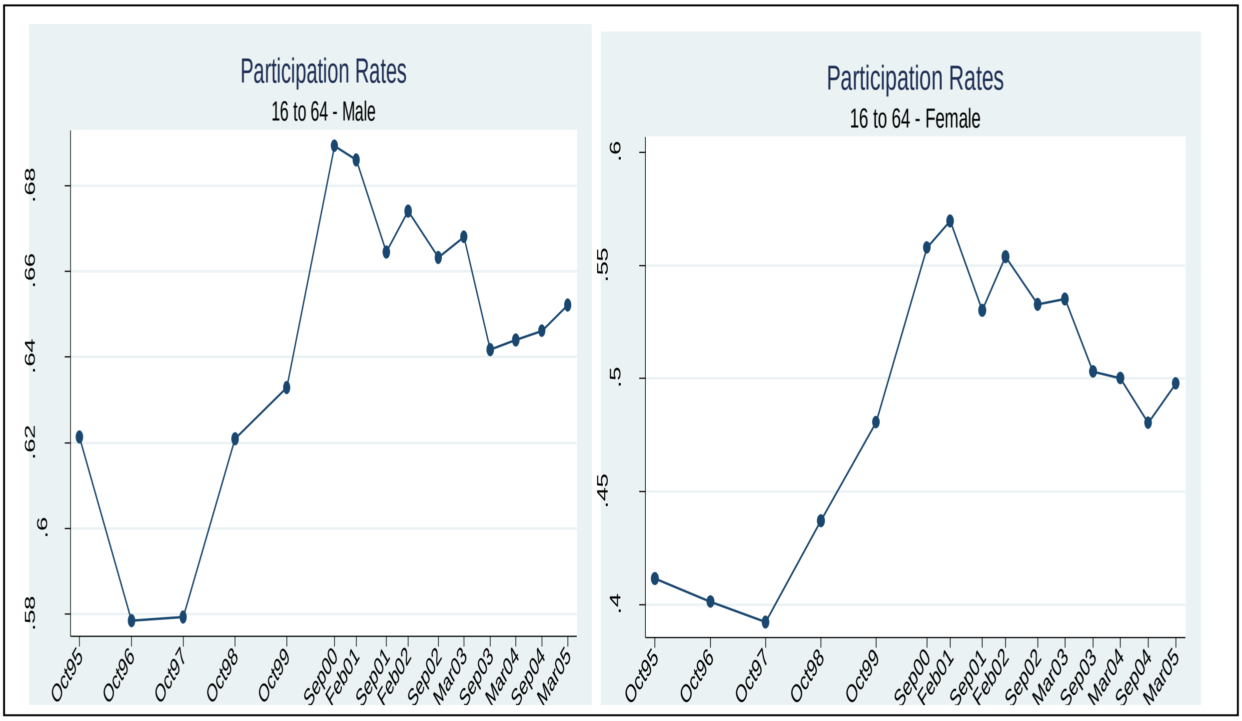
Apart from the change in production methods occasioned by the above-mentioned factor changes, another feature of the South African economy is a change in the structure of its labour market. As noted by Bhorat (2004), the democratic government in South Africa inherited a labour market that was subject to structural shifts and technological change. Between 1970 and 1997, the primary sector of the economy declined by 6.7 percentage points, the secondary sector virtually stagnated (it actually lost about 1.5 percentage points), while the service sector increased by 8.4 percentage points as a share of GDP (Bhorat & Hodge, 1999). A closer look across the sectors shows that there was an increase in the demand for skilled occupational groups. Thus, the general trend was one where a change in production methods (mechanization) led to the displacement of unskilled labour, while a change in the structure of the economy away from primary production to service-oriented production saw an increase in the employment of highly-skilled labour (e.g. professionals) relative to unskilled labour. As shown in Lewis (2002), while the unemployment rate among skilled labour was 16% in 2000, semi-skilled and unskilled labour experienced an unemployment rate of over 50%.

Perhaps, the substantial job loss among the unskilled workforce has led to the view that the South African economy experienced jobless growth early on in the post-apartheid period. This has however been debunked by Bhorat (2004) who observed that the economy created approximately 1.6 million jobs between 1995 and 2002; the economy only witnessed relatively low employment intensity of growth (economic growth was about 2.8% per annum while employment growth was about 2.1% per annum over the period). The caveat however is that part of the increase in the number of the employed between 1997 and 2001 might have been due to better capturing of the informal sector of the economy in household surveys and not necessarily the creation of new jobs (Bhorat, 2004; Casale, 2004).

More recently, Banerjee et al. (2008) observed that though fairly low in the 1970s, the unemployment rate (by the International Labour Organization classification) just before the end of apartheid in 1994 was about 13%. This rose to about 15% and 26.7% in 1995 and 2005 respectively. Using the broad classification of unemployment (i.e. including the discouraged unemployed among the unemployed), Banerjee et al. observed that the unemployment rate rose from 28% in 1995 to 41% in 2005. Narrow and broad unemployment rates for prime-age adults were about 24% and 30% respectively in 2008 (Ranchhod, 2009), and have remained largely so in the current period.

Though the above discussion has explained the causes of high unemployment as emanating from demand-side factors, supply-side factors are equally important. LFP rates exhibited an upward trend between the dawn of democracy and September 2000 (mainly driven by a ten percentage point increase in the female participation rate) and a downward trend from 2000 to 2005[3] (see Figure 2.2), especially in the rural areas (Banerjee et al., 2008). Further declines have been recorded since then (Statistics South Africa, 2012a).

Figure 2.2: Labour force participation trends

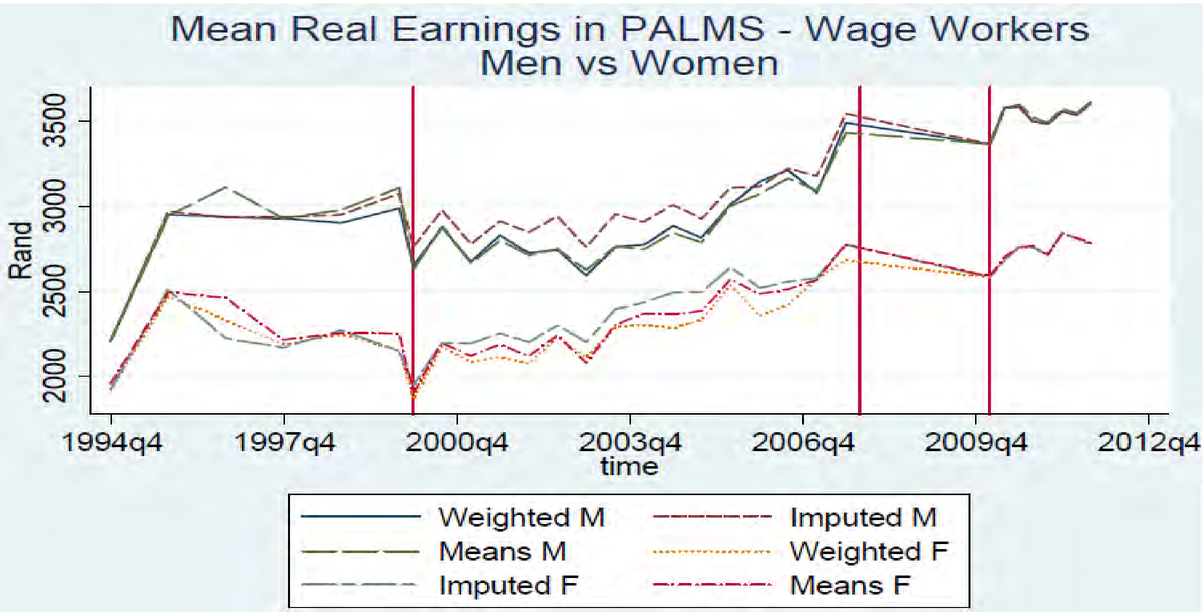


Source: Banerjee et al., 2008; ILO classification

[3] Banerjee at al. (2008) suggested that part of this decline may be due to the incomparability of the October Household Survey and Labour Force Survey (the datasets upon which the analysis was based), a point amply made by Wittenberg (2014).

Another feature of the post-apartheid labour market is the strong and persistent wage-gender gradient. A look at the mean real earnings series for wage workers between 1994 and 2012 (Figure 2.3) shows that male earnings dominated that of females irrespective of the technique used for computation. From the figure, the gender gap apparently did not narrow over time. Similarly, the wage-race gradient for wage workers has always been strong and persistent between whites and other racial groups, though the white-Asian gap appears to be narrowing over time. This finding echoes earlier evidence in South Africa, where average wage among whites differed from that of blacks by a factor of five (Schultz & Mwabu, 1998b). This is unlike what is obtained for the self-employed where the gaps were not substantial though the order of dominance remained as among wage employees, i.e. whites, Asians, coloureds and Africans in descending order (Wittenberg, 2014).

Figure 2.3: Real earnings of wage workers by gender



Source: Wittenberg (2014)

In a similar vein, the South African labour market is characterized by substantial wage inequality. Earlier, it was observed that South Africa is one of the most income unequal countries in the world as evidenced by its high aggregate Gini coefficient (Leibbrandt et al., 2011). LFP, employment and labour absorption patterns have played a substantial role in inducing such inequality, as participation and labour absorption rates were highest among the top income deciles. This resulted in higher employment and lower unemployment rates (and the attendant higher income) among these high income groups despite unemployment rates generally being on an upward trajectory in the country since 1993 (Leibbrandt et al., 2011). Little wonder that labour market earnings accounted for 88.3% and 85% of income inequality in 1993 and 2008 respectively (Leibbrandt et al., 2012). Furthermore, wage inequality differs substantially across industries and occupations. For instance in analysing wage trends and inequality in South Africa, Hlekiso and Mahlo (2006) observed that the 95/50 wage ratio across industries in 2005 was highest in the agriculture, fishery and forestry industry at 8.8 and lowest among private households at 1.5. By occupational groups, the highest ratio of 11.5 was found among skilled agricultural and fishery workers while the lowest ratio of 2.7 existed among clerks.

Related to wage inequalities is wage discrimination. The apartheid system institutionalized race-based discrimination in virtually all facets of life. But with its abolition and the advent of democracy, it has been noted that race-based discrimination has become relatively muted and mainly mentioned in anecdotes (Seekings, 2008). Perhaps, this is the result of legislations like the Promotion of Equality and Prevention of Unfair Discrimination Act, 2000 passed in the wake of apartheid. This notwithstanding, empirical evidence suggests that race continues to be a marker of labour market discrimination in South Africa. For instance, McCord and Bhorat (2003) maintained that while evidence suggests that schooling and experience are among the most important determinants of earnings differentials internationally, earnings differentials in South Africa are mainly driven by racial discrimination and mobility barriers that have to do with urban-rural and formal-informal earnings differentials. Furthermore, gender has also been a source of wage discrimination in the South African labour market (Grün, 2004).

The foregoing has given a brief overview of the South African labour market with special focus on LFP, wage inequality and wage discrimination, key issues to be tackled in this thesis. In the next section, I conduct a descriptive analysis of the South African labour market using the NIDS dataset.

2.4 THE SOUTH AFRICAN LABOUR MARKET (2008-2012): EVIDENCE FROM THE NATIONAL INCOME DYNAMICS STUDY

This section provides an in-depth descriptive analysis of labour market outcomes, especially those critically analysed in the thesis, using the NIDS dataset. The purpose of this analysis is to afford the reader an in-depth understanding of the dataset to be used in this analysis, especially when compared with previous empirical evidence. Some of the statistics presented below have no doubt appeared elsewhere (Cichello et al., 2012; Ranchhod, 2009). But the following discussion, taken as a whole, provides an important context to the broad themes covered in this thesis using the NIDS dataset. Similar descriptive evidence on the interaction between health and key labour market outcomes are presented in Chapter 4-Chapter 6 depending on the particular labour market outcome examined. This is to avoid repetition as well as keeping each of those chapters focused.

2.4.1 Labour supply outcomes and transitions in South Africa (2008-2012)

Every adult who responded to any of the employment questions was classified into one of four mutually exclusive labour market categories: not economically active (NEA), unemployed (discouraged), unemployed (strict/searching), and employed. The NEA are respondents who were not employed and did not want to work at the time of the interview (e.g. students, retirees and home-makers); the discouraged unemployed are respondents who wanted to work but were not working and had not taken any active step to look for employment in the previous four weeks. The strict/searching unemployed are the unemployed who had taken active step(s) toward securing employment in the previous four weeks; while the employed are people who were engaged in some productive activity usually for the purpose of earning money (Ranchhod, 2009). Unless otherwise stated in this thesis, LFP is used broadly to refer to the discouraged unemployed, searching unemployed and the employed. Table 2.1 shows the distribution of labour market categories in wave 1 and wave 3 as well as transitions across time.

Table 2.1: Labour market transition (2008-2012)

			Labour market status (2012)			
			NEA	Discouraged	Searching	Employed
	%		24.2	3.4	15.7	56.7
Labour market status (2008)	21.2	NEA	45.8	4.6	19.5	30.1
	6.0	Discouraged	30.7	9.7	28.6	31.1
	17.6	Searching	25.9	2.9	27.9	43.3
	55.3	Employed	16.1	2.2	9.2	72.5

Sample is restricted to adults aged 20-56 years in wave 1 who were interviewed in both waves; sample is nationally representative and corrected for non-random attrition using appropriate weights
Source: Own calculations

As the table indicates, the proportion of the working age population (as defined in this study) who were non-economically active and the employed increased by three percentage points and one percentage point respectively while the proportion of the discouraged and searching unemployed both declined over the four-year period. The most stable labour market category was the employed group (72.5% of the employed in wave 1 remained employed in wave 3), followed by the NEA. On the other hand, the least stable group was the discouraged unemployed. Furthermore, about a third of the discouraged unemployed as well as a quarter of those actively searching in wave 1 dropped into economic inactivity in wave 3. This pattern of transitions no doubt contributed to declining unemployment rates as will be seen subsequently.

A gendered analysis of labour market outcomes in both waves (available on request) indicates that women have not fared well in the labour market relative to men. This finding also corroborates previous evidence (Casale, 2004). More women were stuck in economic inactivity and discouragement than men in proportional terms across waves. The proportion of women in the working age population who were NEA was about double that of men in each wave, while the employed proportion among women was about two-thirds that of men in both waves. Unemployment was also more concentrated among women. Furthermore, gender-related transitions over these years appeared to make women worse off relative to men. For instance, while half of women who were not economically active in wave 1 remained so in wave 3, it was only a third for men. Among women, only 27% of the wave 1 non-economically active became employed in wave 3 while the figure was 37% for men. Moreover, while only 66% of employed women in wave 1 were still employed in wave 3, 78% of men retained their "employed" status.

It may be argued that the above gender-related outcomes are due to greater aversion for work among women relative to men. But as shown in Table 2.4 below, the proportion of economically inactive men who either disliked work in general, or did not like available jobs exceeded that of women especially in wave 3. Therefore, the above differences are not likely to have been caused by a taste against work among women. Thus, there appears to be evidence of gender-biased labour market participation in favour of men, a finding that corroborates earlier evidence in South Africa (Klasen & Woolard, 1999). Added to the fact that the proportion of household heads made up of women increased across waves (in fact about 55% of households in wave 3 were headed by women), such gender-biased labour market outcomes present a worrying picture for household welfare as households with employed heads were significantly ($p<0.01$) richer on average than those where the head was not employed, in each wave. Average monthly household nominal income in households with an employed (unemployed) head was R7995 (R3546) in wave 1 and R11709 (R6244) in wave 3.

A racial decomposition of labour market status over the waves in Table 2.2 shows that whites had the highest LFP and lowest unemployment rates in wave 3 irrespective of definition (Africans had the lowest LFP and highest unemployment rates in wave 1). Also, aggregate strict (broad) unemployment rate fell by three (five) percentage points, from 24.1% (29.9%) to 21.6% (25.2%). However, caution should be exercised in celebrating this decline as it was partly due to declining LFP (especially broad LFP); aggregate strict (broad) LFP rate fell from 72.9% (78.9%) to 72.4% (75.8%). Also, the racial differences in unemployment were substantial. For instance, the African unemployment rate was about three times that of whites in wave 1 irrespective of definition. It was 4-5 times in wave 3. Evidently, race seems to play a vital role in determining LFP and unemployment in South Africa. This is confirmed by Wittenberg (2014).

Table 2.2: Racial distribution of labour market categories

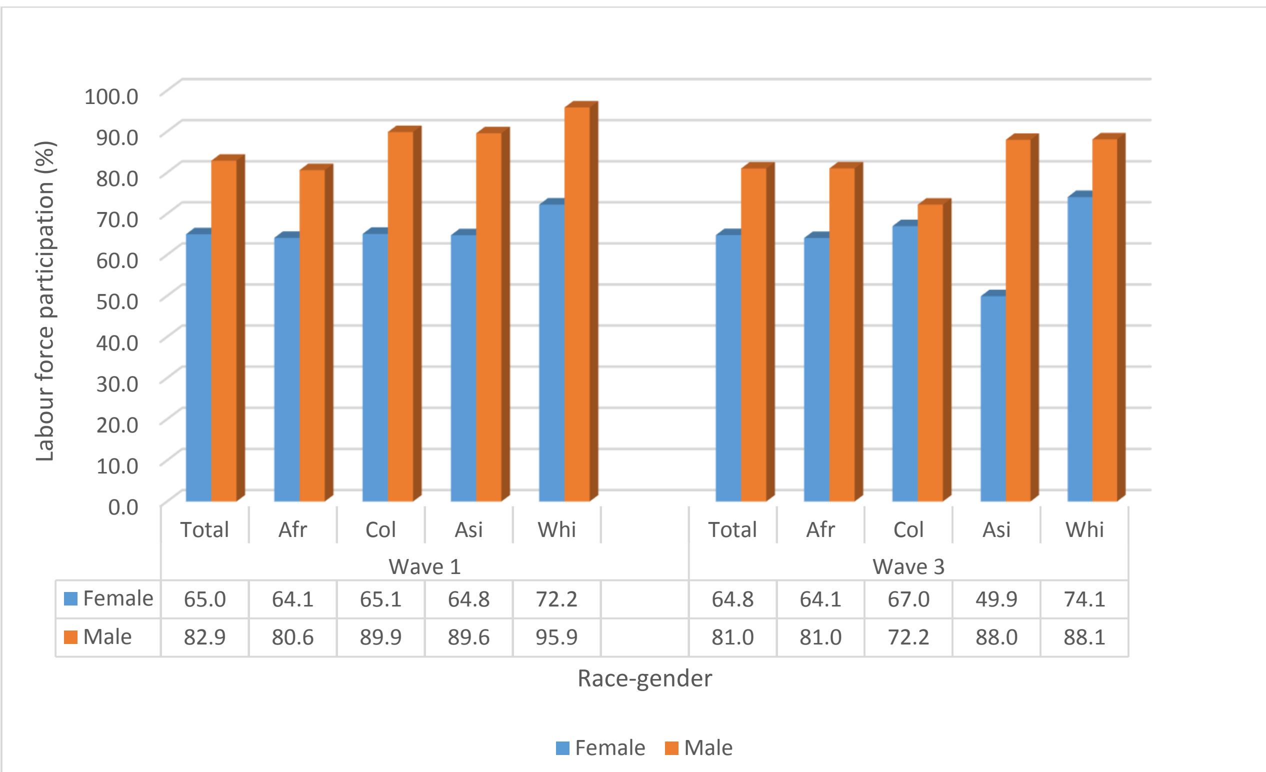
			Percentage in each category in wave 3								**% point change (wave 3-wave 1)**			
			i	ii	iii	iv		v	vi	vii	viii	ix	x	xi
				Unemployed		Employed	LFP rate		Unemployment rate		LFP rate		Unemployment rate	
		N	NEA	Disc†	Searching		Strict	Broad	Strict	Broad	Strict	Broad	Strict	Broad
Racial categories	Aggregate	9229	24.2	3.4	15.7	56.7	72.4	75.8	21.6	25.2	-1.0	-3.0	-3.0	-5.0
	African	7253	24.7	3.4	18.0	54.0	72.0	75.4	25.0	28.3	0.0	-2.0	-3.0	-5.0
	Coloured	1506	25.9	4.6	9.8	59.7	69.5	74.1	14.1	19.4	-6.0	-8.0	-2.0	-4.0
	Asian	134	27.4	4.7	4.4	63.5	67.9	72.6	6.5	12.5	-8.0	-9.0	0.0	-1.0
	White	336	16.9	2.3	3.6	77.2	80.8	83.1	4.4	7.0	-1.0	-2.0	-4.0	-4.0

a) Sample is restricted to adults aged 20-56 years in wave 1 who were interviewed in both waves.
b) All wave 1 (wave 3) proportions have been weighted using post-stratification (panel) weights.
c) Conventional definitions were used in calculating LFP and unemployment rates: $Strict\ LFP\ rate = \frac{(iii+iv)}{(i+ii+iii+iv)} X100; Broad\ LFP\ rate = \frac{(ii+iii+iv)}{(i+ii+iii+iv)} X100;\ Strict\ unemployment\ rate = \frac{(iii)}{(iii+iv)} X100; Broad\ unemployment\ rate = \frac{(ii+iii)}{(ii+iii+iv)} X100$ (d) (i)-(iv) sum to 100; † Discouraged unemployed

Source: Own calculations

Figure 2.4 further disaggregates race-based LFP rates by gender in both waves. Most gender-race groups experienced a decline in LFP rate across waves. Looking at the male sub-sample, all but the African group experienced LFP decline across waves. This feature was most pronounced among coloured males, with an 18 percentage point decline. A consistent trend however is that male LFP was substantially higher than female LFP in both waves across all racial groups. An important caveat when considering Figure 2.4 (especially the very low female Asian LFP rate in wave 3) is that some of the cell sizes are very small. For instance, the number of female Asian and white respondents in the sample who were discouraged and searching for jobs in wave 3 was very small.

Figure 2.4: Distribution of race-gender strict labour force participation (wave 1 & wave 3)



Statistics appropriately weighted with post-stratification and panel weights

Source: Own calculations

Table 2.3 below analyses key labour market outcomes in wave 3 as well as changes in those outcomes between wave 1 and wave 3 by gender, age and marital status. About a third of women were economically inactive in wave 3 while less than one-fifth of men belonged to this category. Moreover, while strict male LFP was about 81%, that of females was only 65%. Also, strict female unemployment rate was nine percentage points higher than that of males. Neither gender recorded an increase in LFP or unemployment rate.

A more careful analysis of these gender-related LFP (see Table 2.4) and unemployment patterns shows that the most common reasons why inactive women did not search for jobs in wave 1 were looking after children (20.5%) and sickness (20.3%). Only 1.4% and 3.2% of women reported not searching due to work aversion and not liking available jobs respectively. For men however, being a full time student (33.1%) and sickness (29.6%) were the most common reasons. Child care was also the most common cause of female economic inactivity in wave 3. These findings reveal at least two important facts: child care is one of the most important impediments of female participation in the South African labour market; and ill-health exerts great cost on the economy by preventing economic agents from actively participating in the labour market. Evidence from the US also corroborates the deleterious effect of child care on female LFP (Anderson & Levine, 1999). It is also interesting that a high reservation wage relative to the offered wage does not appear to be a major impediment of (especially female) LFP as only 0.4% (2.3%) of women (men) cited it as a reason for non-participation in wave 1 while the wave 3 figures were 1.6% (7.9%). This confirms prior evidence that high reservation wage is not a major cause of non-participation in the South African labour market (Banerjee et al., 2008; Nattrass & Walker, 2005). In another analysis (not shown in Table 2.4 but available on request), family commitments (33%) was the second most common reason why unemployed economically active women did not accept employment offers in wave 1, and the most common reason in wave 3. However, it was among the least cited reasons for men (1% and 0% in wave 1 and 3 respectively). But caution should be exercised in interpreting these employment refusal figures because of very small cell sizes. Thus, there are arguably gender-based differences in the role of family commitments on participation in the labour market. It is however surprising that a nontrivial proportion of each gender perceived themselves to be too old to participate in the labour market given the working age population chosen in this study.

Given the very high levels of youth unemployment in South Africa, it is important to ascertain patterns of LFP and unemployment by age. This is also shown in Table 2.3. The age groups as shown in the table are in relation to wave 1 age classification; therefore, wave 3 age groups are roughly four years more than wave 1 age groups. In wave 3, economic inactivity continuously declined from 24-29 years to 34-39 years, and rose continuously from 40 to 60 years. This pattern is reflected in the LFP pattern across different age groups. However, unemployment was most acute among the youngest age cohort as expected and was continuously decreasing as respondents got older. For instance, strict unemployment rate among the youngest (oldest) cohort was 14.2 (10.6) percentage points higher (lower) than the population average.

Table 2.3 also indicates substantial differences in the incidence of unemployment across marital status groups. The classification of marital status groups here was informed by the need to separate those currently living together as couples from the currently “single”. That is, respondents still married or cohabiting were classified as “married” while those who had never married, the divorced/separated and widows/widowers were classified as “not married”. Thus, “married” and “not married” are simply terms used to denote whether or not people were currently living together as couples and does not have any legal connotation. Such a distinction is important especially in light of the theory of marriage and labour market participation which suggests that single women are more likely to be economically active compared to their married counterparts while the converse holds for men (Becker, 1981). This might explain why a larger proportion of the not married group were searching for jobs relative to the married in both waves. Though LFP rates were almost identical between both groups in wave 3, unemployment rates were substantially lower among the married relative to the non-married in both waves: strict (broad) unemployment rate in the married group was 14.2% (17.2%) and that of the non-married group was quite higher at 26.4% (28.8%). These differences were mainly driven by higher (lower) proportion of searching unemployed (employed) in the non-married group relative to married respondents.

Though not supported by all scholars, the dominant literature on family specialization suggests that single women are more likely to be economically active than their married counterparts while the converse holds for men (Becker, 1981; Oppenheimer, 1997). Therefore, labour market participation of respondents belonging to different marital status groups is expected to differ across gender. For women (table available on request), I found that though LFP rates were largely similar between married and non-married respondents, unemployment rates among non-married women were substantially higher than among their married/cohabiting counterparts (about ten percentage points higher for both strict and broad unemployment rates in both waves). This is apparently a contradiction of the Becker theory. This was largely driven by a combination of higher (lower) proportion of searching (employed) non-married women relative to their married/cohabiting counterparts as in the aggregate case. Perhaps, this captures the effects of networks if married/cohabiting women have more job-related networks (and therefore more labour market-related information and connection) than the non-married due to say, employed husbands/partners. This is likely, given that married/cohabiting men had far lower unemployment rate compared to their non-married counterparts on the average (strict and broad unemployment rates among non-married men were about three times that of the married/cohabiting in wave 1 while the ratio only declined to approximately 2.5 times in wave 3; results are available on request).

In a related development, prior evidence linked declining marriage rates in South Africa in the mid to late 1990s to increased female (especially African women's) LFP then (Casale & Posel, 2002). However, evidence from NIDS shows increased marriage rates between 2008 and 2012 across all racial and gender groups (table available on request). This might be a contributory factor to non-increasing female LFP in this period.

Table 2.3: Distribution of gender/age/marital status across labour market categories in wave 3

		Percentage in each category in wave 3								**% point change (wave 3-wave 1)**			
		i	ii	iii	iv	v	vi	vii	viii	ix	x	xi	xii
			Unemployed		Employed	LFP rate		Unemployment rate		LFP rate		Unemployment rate	
	N	NEA	Disc	Searching		Strict	Broad	Strict	Broad	Strict	Broad	Strict	Broad
						Gender							
Aggregate	9203	24.3	3.3	15.7	56.7	72.5	75.7	21.7	25.1	0.0	-3.0	-2.0	-5.0
Female	5425	31.7	3.5	17.1	47.6	64.8	68.3	26.5	30.3	0.0	-5.0	-5.0	-8.0
Male	3778	16.1	2.9	14.1	66.9	81.0	83.9	17.5	20.3	-2.0	-3.0	0.0	0.0
						Age categories*							
Aggregate	9229	24.2	3.4	15.7	56.7	72.4	75.8	21.6	25.2	-1.0	-3.0	-3.0	-5.0
20-25	2274	20.9	3.5	27.1	48.5	75.5	79.1	35.8	38.7	15.0	10.0	-4.0	-8.0
26-30	1331	18.6	3.9	16.8	60.8	77.6	81.4	21.6	25.3	-2.0	-5.0	-10.0	-11.0
31-35	1248	16.6	2.5	16.0	64.9	80.9	83.4	19.8	22.2	2.0	-1.0	-3.0	-6.0
36-40	1119	20.2	4.1	13.1	62.6	75.7	79.8	17.3	21.5	-4.0	-7.0	1.0	-1.0
41-45	1151	25.6	3.3	10.3	60.8	71.1	74.4	14.5	18.2	-9.0	-9.0	-3.0	-3.0
46-50	1056	31.6	5.0	7.7	55.7	63.4	68.4	12.2	18.6	-8.0	-8.0	-4.0	-3.0
51-56	1050	48.9	1.9	5.4	43.8	49.2	51.1	11.0	14.3	-12.0	-14.0	1.0	0.0
						Marital status**							
Aggregate	8190	25.8	2.5	15.7	56.1	71.8	74.3	21.8	24.4	-1.0	-5.0	-2.0	-6.0
Not married	5318	25.6	2.4	19.0	53.0	72.0	74.4	26.4	28.8	1.0	-3.0	-2.0	-6.0
Married	2872	26.1	2.5	10.2	61.3	71.4	74.0	14.2	17.2	-5.0	-7.0	-2.0	-3.0

(a) Sample is restricted to adults aged 20-56 years in wave 1 who were interviewed in both waves. (b) All wave 1 (wave 3) proportions have been weighted using post-stratification (panel) weights.(c) Conventional definitions were used in calculating LFP and unemployment rates: $strict\ LFP\ rate = \frac{(iii+iv)}{(i+ii+iii+iv)} X100$; $Broad\ LFP\ rate = \frac{(ii+iii+iv)}{(i+ii+iii+iv)} X100$; $Strict\ unemployment\ rate = \frac{(iii)}{(iii+iv)} X100$; $Broad\ unemployment\ rate = \frac{(ii+iii)}{(ii+iii+iv)} X100$. (d) (i)-(iv) sum to 100. (f) *Age categories are in reference to wave 1 categories. This implies wave 3 age groups are roughly four years more than wave 1 age groups. (e) **Not married→ never married, divorced/separated, or widow/widower; Married→ married, or living with partner

Source: Own calculations

Table 2.4: Reasons for non-participation by gender in percentage (All column percentages sum to 100)

	Wave 1 (%)		**Wave 3 (%)**	
	Female	Male	Female	Male
Too old	11.8	7.6	20.4	18.9
Full time student	12.7	33.1	3.5	5.1
Sick/disabled	20.3	29.6	17.7	32.8
Don't like available jobs	3.2	2.2	3.6	7.1
Don't like working	1.4	3.0	2.8	3.5
Domestic duties	11.0	2.0	14.3	0.7
Looking after children	20.5	0.4	21.5	0.5
Job search is expensive	3.4	8.9	6.9	18.4
Low wages	0.4	2.3	1.6	7.9
Cooking/cleaning	4.9	0.6	3.4	0.1
Other	6.6	6.8	2.4	5.1
Pregnant	2.0	-	1.8	-
Still searching	1.9	3.6	-	-
Total sample size	**1861**	**700**	**1858**	**615**

a) Sample is restricted to adults aged 20-56 years in wave 1 who were interviewed in both waves.

b) All wave 1 (wave 3) proportions have been weighted using post-stratification (panel) weights.

Source: Own calculations

One lasting legacy of apartheid and an arguably important determinant of labour market outcomes in South Africa is the spatial distribution of the population. Apartheid-era racial segregation led to the spatial restriction of non-white population groups (Rospabe & Selod, 2006). As a result, there is disproportionate representation of different population groups in different locations across the country. For instance, whites are generally over-represented in rich urban neighbourhoods while Africans mostly populate the rural and poor informal urban locations. This pattern of living arrangements is reflected in the NIDS dataset. For instance, Africans ((who made up about 79% of the population in 2011 (Statistics South Africa, 2012a)) constituted 99.9% of traditional authority dwellers and 96.3% of urban informal areas in wave 3. On the other hand, whites (who accounted for about 9% of the population) constituted 14.4% of urban formal areas and 0% of traditional authority areas in wave 3.

One thing that is apparent is the wide geographical variation in the level of economic activity and unemployment incidence in South Africa even within the same racial group. For instance, the African broad LFP rate ranged from 65% in traditional authority areas to 87% in rural formal areas in wave 1. On the other hand, the wave 1 African broad unemployment rate ranged from 16% in rural formal areas to 43% in traditional authority and urban informal areas. African unemployment was also most severe in traditional authority areas in wave 3. Though participation rates were not starkly different between Africans, coloureds and whites in urban formal areas in wave 1, unemployment incidence varied widely; African and coloured broad unemployment rates were 28% and 24% respectively while that of whites was 11%. A similar pattern obtained in wave 3. The foregoing shows that one's location appears to be an important determinant of LFP and unemployment, while reinforcing the racial undertones of unemployment in South Africa. Similar findings have been made in previous studies (Dinkelman & Pirouz, 2002; Ntuli & Wittenberg, 2013).

Related to location in a South African context is job search. Given that apartheid-era structures resulted in jobs being mainly situated in formal cities, it is expected that job search will be more expensive for job seekers living away from cities (Banerjee et al., 2008). For the currently employed, the average number of months spent in unemployment before securing the current job was 39.7 months in wave 1. But this duration varied substantially across locations. It was highest in traditional authority areas (53 months) and lowest in rural formal areas (35 months). Wave 3 figures are unavailable as this question was dropped from the questionnaire. For the searching unemployed, average transport cost spent searching for jobs in the past week was R103 in wave 1. Disaggregating this amount across locations, job search was most expensive in rural formal areas (R381) and cheapest in urban informal areas (R69). In wave 3, the overall average increased to R137. It was costliest in traditional authority areas (R145) - only slightly more expensive than in rural informal areas (R139) - and also cheapest in urban informal areas (R121). In general, residents of urban formal locations incurred the second lowest cost of job search in both waves (R86 and R134 respectively). This is largely supportive of the above thesis of relatively low search costs in cities. These amounts are however substantial especially when viewed from the perspective of an unemployed economic agent. This may be one of the reasons why a large proportion of job seekers subsequently dropped into economic inactivity as earlier shown in Table 2.1.

2.4.2 Wage determination and wage-related transitions in South Africa (2008-2012)

Another issue investigated in this thesis is wage determination. As in the foregoing section, I conduct descriptive analysis of the relationship between wage and key covariates and compare these to other findings as well as examine temporal changes in the relationships. "Wage" as used here, refers to individual real monthly take home pay from main job (casual and self-employment were excluded). Average and median real wages increased from R5491.50 and R2871.80 respectively in wave 1 to R5861.53 and R3256.58 in wave 3 respectively. Thus, there was evidence of real growth in both average and median wages as well as positively skewed wage distributions across waves.

Education is a common measure of human capital accumulation. Respondents with at least a matric[4] qualification earned about three times their less educated counterparts on the average in both waves ($p < 0.01$). Given the high degree of aggregation inherent in this kind of classification, I also disaggregated the analysis by actual years of schooling. Average monthly wage among those with no education was R1529 (N=1130) and R1931 (N=884) in wave 1 and wave 3 respectively. Table 2.5 below shows differences in average monthly wage for each year of schooling relative to no education in both waves.

[4] 'Matric' is a colloquial term used in South Africa to refer to the National Senior Certificate, which is obtained upon the successful completion of the National Certificate Examinations. It is equivalent to roughly twelve years of schooling. A matriculation endorsement or exemption is the minimum requirement for admission to a bachelor's degree programme in South African universities.

Table 2.5: Monthly real wage differentials between various levels of education and no education (Rands)

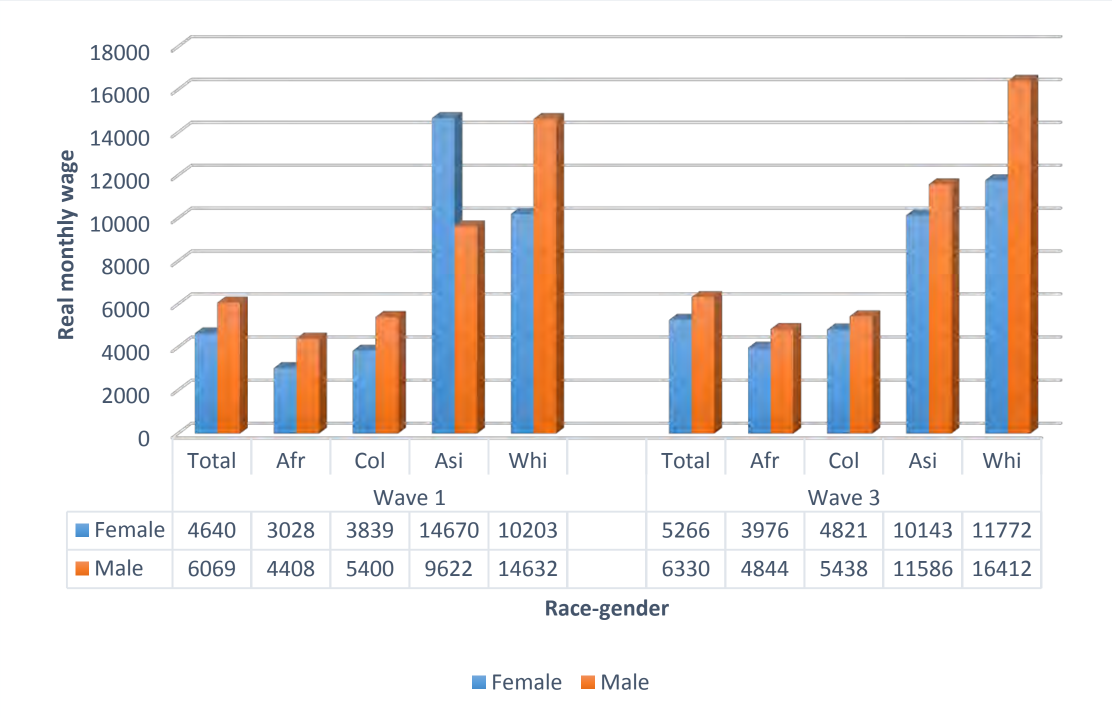
Years of schooling	Wave 1		Wave 3	
	N	Wage differences	N	Wage differences
1	70	-511.04**	52	6,030.21
		(0.02)		(0.14)
2	149	-31.61	132	431.28
		(0.88)		(0.11)
3	289	596.92	247	342.63
		(0.13)		(0.25)
4	381	884.14	350	161.13
		(0.13)		(0.63)
5	458	325.79	360	392.57
		(0.10)		(0.19)
6	581	470.07**	451	458.25
		(0.02)		(0.23)
7	885	749.06***	736	513.83*
		(0.00)		(0.08)
8	861	693.64***	706	1,319.54***
		(0.00)		(0.00)
9	949	788.35***	772	858.28***
		(0.00)		(0.01)
10	1327	2,075.26***	1054	1,595.36***
		(0.00)		(0.00)
11	1539	1,744.06***	1407	1,506.42***
		(0.00)		(0.00)
12	2461	4,492.29***	1769	4,003.30***
		(0.00)		(0.00)
13	860	7,988.85***	1127	6,267.88***
		(0.00)		(0.00)
15	143	10,763.47***	138	15,763.75***
		(0.00)		(0.00)
16	121	12,010.89***	158	12,438.85***
		(0.00)		(0.00)
18	43	27,213.01***	45	26,665.31***
		(0.00)		(0.00)

p values in parentheses; *** p<0.01, ** p<0.05, * p<0.1; estimates corrected for complex survey design and non-random attrition

Generally, there was no significant earnings differential between respondents without any education and those with up to five years of schooling in wave 1 (indeed, individuals with only a year of schooling earned significantly less than those without any education). However, from the sixth year of education onwards, statistically significant positive differences emerged up to the highest number of years of schooling (i.e. eighteen). These differences increased with years of schooling. The differences ranged from R470 to R27213 per month. Even more significant is the large returns associated with at least one extra year of schooling above the matric level compared with quitting school at matric. For instance, the thirteenth year of schooling was associated with an extra R3500 per month on the average compared to stopping school at matric level, while eighteen years of education (equivalent to having at least a master's degree) was associated with an extra R23000 per month on the average relative to a matric. A similar trend occurred in wave 3 (here, the returns to schooling only kicked in at the seventh year). Lam (1999) also found positive and increasing returns to schooling at each year of schooling from the fourth year onwards in both South Africa and Brazil (it was for all schooling years in Brazil). In addition to South Africa, this pattern has also been confirmed in Nigeria, Ghana, Kenya and Ivory Coast (Schultz, 2004).

Earlier, I showed that gender played an important role in determining LFP and unemployment. I hereby examine how important it is in wage determination once one secures employment. Real average monthly wage for females (males) was R4640 (R6069) in wave 1 as shown in Figure 2.5 below. This implies that male wage in wave 1 was 31% higher than female wage on the average. A racial analysis of this gap (see Figure 2.5) shows that male wages exceeded female wages by 41-46% across all racial groups (with the biggest difference occurring among Africans) except the Asian group where average female wage exceeded male wage by 52%, though not statistically significant at conventional levels. But as earlier noted, Asian statistics are to be viewed with caution given small sample size. The aggregate gender wage gap narrowed to 20% in wave 3, while only the African and white gaps (21% and 39% respectively) were statistically significant ($p<0.05$). Thus, while the African gap declined from 46% to 21% between 2008 and 2012, the white gap only fell from 43% to 39% over the same period. The substantial percentage drop in the African gap was necessitated by a 31% rise in African women's wage earnings compared to only 10% for African men. On the other hand, wage growth for white women and men were similar: 15% and 12% respectively. Nontrivial gender-based wage gaps (in favour of men) has been shown to be a consistent feature of the post-apartheid South African labour market for both employees and employers/the self-employed (Wittenberg, 2014).

Figure 2.5: Race-gender wage distribution (2008-2012)



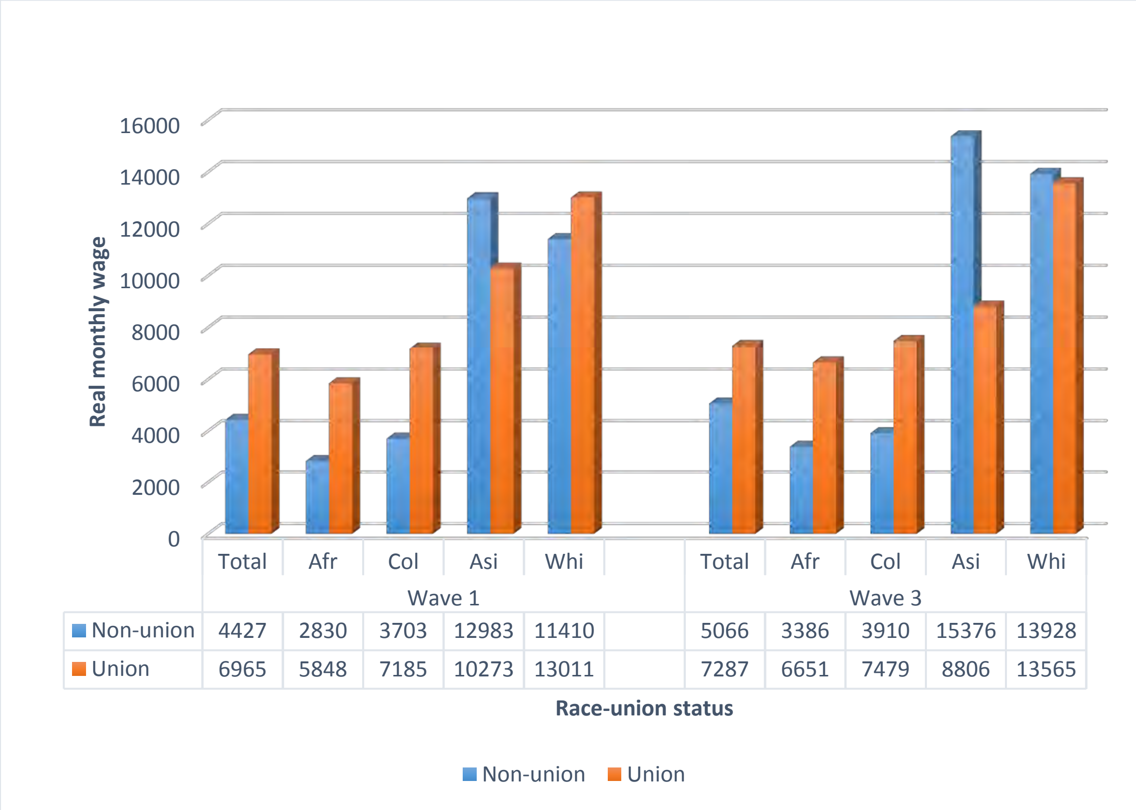
Statistics appropriately weighted with post-stratification and panel weights

Source: Own calculations

Unionization remains a vexing issue in South Africa. Proponents argue that unions ensure more egalitarian wage distribution in firms (Freeman, 1982). However, others maintain that unionization increases unemployment, the result of wages rising faster than productivity (Banerjee et al., 2008). Overall in the dataset analysed, unionization was associated with higher real monthly wage on average in both waves (Figure 2.6). In wave 1, average monthly wage of trade union members (R6965.09) was 57% higher than that of non-union members (R4426.75) while it was higher by 44% in wave 3. These wage differences were statistically significant in each wave ($p<0.01$). Though the average wage of union members only grew by 4.6% between both waves while that of non-members grew by 14.4%, it is important to point out that non-members' average wage in wave 3 (i.e. R5065.74) was still less than union members' average wage in wave 1 (R6965.09). These huge wage differentials may also encompass selection issues. For instance, it has been argued that union members are generally better skilled than non-members in South Africa (Seekings, 2014). Controlling for education ("at least matric" vs. "no matric"), I found that the monthly average wage of unionized workers without a matric qualification (R3995.85) was still significantly higher than the corresponding non-members' average wage (R2068.24). Thus, education-related selection into union membership is not wholly responsible for such high union premium. The difference was however not statistically significant (even at the 10% level) for those with at least a matric qualification. These trends were replicated in wave 3.

Some studies have shown that the union premium in South Africa is not race-neutral (Chamberlain & Van der Berg, 2002; Schultz & Mwabu, 1998b). A similar pattern holds here as union premium among Africans, coloureds and Asians were statistically significant in wave 1 at conventional levels, while that of whites was not. In wave 3, only those of Africans and Asians were significant. Generally, unionization appears to favourably affect Africans; among Africans, union members' average wage was more than double that of non-members in wave 1. It only declined to 75% in wave 3 (see Figure 2.6). A very high African union premium, especially among low-earners, is not new in South Africa; Schultz and Mwabu (1998b) found that average African male union members' wage in the bottom decile of the wage distribution was 145% of their non-union counterparts in 1993. On the other hand, I found that among whites, the average wage of union members was only 14% higher than that of non-members in wave 1 while union members actually suffered a wage penalty of 3% compared to non-union members in wave 3 (though not statistically significant). In a related development, Schultz and Mwabu (1998b) found a 24% conditional wage penalty among white union members in the 90th percentile of the wage distribution.

Figure 2.6: Race-union status wage distribution (2008-2012)



Statistics appropriately weighted with post-stratification and panel weights
Source: Own calculations

Furthermore, wage differentials may arise due to differences in occupational categories. Respondents' primary jobs were grouped into three occupational classes: managerial/professional (legislators, senior officials, managers, professionals, technicians and associate professionals); semi-skilled (clerks, service workers, shop and market sales workers, skilled agricultural and fishery workers, craft and related trades workers as well as plant and machinery operators and assemblers), and the elementary occupations. This classification is in line with Cichello et al. (2012). These groups are no doubt highly aggregated but provide an indication of occupational differentiation nonetheless.

There were numerically substantial and statistically significant ($p<0.01$) wage differentials between these occupational classes. For instance in wave 1, the real monthly wage of the managerial/professional group (R10315.70) was six times that of the elementary group (R1717.32) and two and a half times that of the semi-skilled category (R4182.55) on average (see column 3 of Table 2.6). These proportions declined somewhat in wave 3 but the differences still remained nontrivial (column 6 of Table 2.6). It is important to observe that though these occupational classes are expected to mirror educational and skill differences between workers, there may be underlying non-skill factors driving some of these earnings differentials. For instance, workers in the managerial/professional group with no matric still earned more than twice ($p<0.01$) that of their counterparts in the elementary category and almost one and half times ($p<0.05$) that of their semi-skilled counterparts in wave 1. These differences persisted into wave 3. Moreover, the average wage for respondents with at least a matric exceeded that of non-matric holders within each occupational group ($p<0.01$). These wage differences are shown in Table 2.6 below. Also, regarding the manner of transitions across these occupational categories, the most immobile category was the elementary occupations as 81% remained in this category over the four-year period, while the most mobile was the semi-skilled as only 65% remained in this category in both waves. About 18% of the semi-skilled category in wave 1 "adversely" moved into elementary occupations in wave 3 (transition matrix available on request).

Table 2.6: Wage distribution across occupational class and education

	(1)	(2)	(3)	(4)	(5)	(6)
		Wave 1			Wave 3	
Occupational category	Matric	No matric	**Total**	Matric	No matric	**Total**
Manager/professional	11517.69	3994.16	**10315.70**	12822.52	4544.34	**11360.49**
Semi-skilled	5787.21	2850.72	**4182.55**	6020.728	3351.90	**4620.199**
Elementary	2514.95	1587.03	**1717.32**	3942.448	2004.22	**2387.675**

Source: Own computation; estimates corrected for complex survey design and non-random attrition

As earlier indicated, average and median real monthly wage increased between waves 1 and 3. In reality, this may not reflect welfare improvement if it was only brought about by more hours of work. It is important to note the existence of measurement error in the weekly hours variable as some weekly hours exceeded 168 (maximum possible hours in a week). Therefore, I set weekly work hours exceeding 105 (i.e. 15 hours/day for 7 days) to missing. Average weekly hours virtually remained unchanged across waves (41.9 hours and 41.7 hours respectively) while the median remained unchanged at 42 hours per week in each wave (not shown in Table 2.7). Therefore, growth in monthly wage was not due to higher working hours in wave 3 relative to wave 1.

Table 2.7: Distribution of average real wage (monthly) and hours (weekly) across industries

	Wave1				**Wave 3**			
Industry	Wage	N	Hours	N	Wage	N	Hours	N
Private households & exterritorial org.	1254.8	376	36.3	336	1421.9	316	35.9	305
Agriculture	1483.2	492	48.0	453	1950.2	410	44.4	401
Mining and quarrying	9019.2	150	44.8	129	8257.6	141	44.8	140
Manufacturing	4696.9	475	44.2	434	5157.6	265	44.4	260
Electricity gas and water supply	6666.6	23	44.1	23	8289.8	47	41.3	46
Construction	3225.0	174	41.7	159	3438.0	123	45.0	119
Wholesale and retail trade	4178.9	433	42.9	400	4805.1	526	40.4	517
Transport, storage and communication	5643.9	112	40.8	102	5084.7	153	45.6	147
Financial intermediation	7478.8	241	45.3	224	7464.4	200	44.6	199
Community, social and personal service	7308.4	776	37.9	735	8420.1	925	39.9	913
Total	**5239.5**		**41.9**		**5917.3**		**41.7**	

(a) Sample is restricted to adults aged 20-56 years in wave 1 who were interviewed in both waves.

(b) All wave 1 (wave 3) proportions have been weighted using post-stratification (panel) weights.

Source: Own calculations

Table 2.7 depicts the distribution of average wages and weekly work hours across industries. Monthly wages were highest in the mining sector in wave 1. This was followed by financial services. The mining industry was second highest in wave 3 behind community, social and personal services in wave 3 (I did not analyse the electricity industry given very small sample size). On the other hand, the private household and agriculture industries earned the lowest average wage in both waves. Table 2.7 reveals deep inequalities across industrial groups in the South African labour market. For instance, average wage in the mining industry was six times that of the agricultural industry in wave 1. This only declined to four times in wave 3. Real wages only declined noticeably in the mining and transportation industries while the other industries (except financial intermediation) experienced average growth in monthly wages.

Average weekly work hours were highest in the agriculture industry in wave 1 and the transport industry in wave 3 but lowest in the private & exterritorial industrial sector in both waves. However, hours worked was similar across industries especially in wave 3 and is therefore not a significant driver of these wage disparities. Overall, most workers in these industries worked full time (defined here as working at least 30 hours per week). As in Klasen and Woolard, (1999), part-time work was more prevalent among women than men: 30% of women and 17% of men worked part-time in wave 1. These declined to 23% and 13% respectively in wave 3.

2.5 CONCLUSION

This chapter has provided an overview of the South African economy and a historical overview of the state of the South African labour market over the post-apartheid period, especially with regard to labour market participation and wage determination. Finally, using the dataset to be analysed throughout this thesis, I conducted a descriptive analysis of labour market participation and wage determination between 2008 and 2012 focusing on factors often considered to be important in labour market models. The results generally conformed to expectations with regard to the relationship between the labour market outcomes of interest and these factors.

Though LFP increased in the mid-90s to early 2000s (mainly due to the influx of African women into the labour force), it generally declined from then onwards. Also, a substantial proportion of the searching unemployed in 2008 moved into economic inactivity by 2012, a trend that is more pronounced among women. This might not be unconnected with substantial search costs faced by the unemployed. I also showed that non-participation is not mainly driven by high reservation wages but by family commitments (for women), full time study (for men), perceived old age and sickness. Furthermore, the analysis showed that labour market differences in outcomes like LFP, unemployment and wages are largely persistent along gender, racial and occupational lines over the post-apartheid period. Also, union membership is associated with higher remuneration especially among Africans, while location plays an important role both in the participation decision and remuneration. A positive finding though, is that many of these gaps reduced between 2008 and 2012. Unlike prior evidence however, marriage rates increased between 2008 and 2012. Also, returns to schooling are very large but only kick in around the sixth or seventh year of formal education.

As shown in the foregoing discussion, ill-health seems to be one of the biggest impediments to labour market participation in South Africa. Consequently, it is important to ascertain the potential role health plays in determining participation. This is the crux of Chapter 4. Also, as subsequently shown, employees in relatively poor health receive significantly lower pay compared to their counterparts who enjoy better health conditions. This issue, as well as health-related differentials in returns to observable workers' characteristics (popularly dubbed wage discrimination) is given in-depth attention in Chapter 5 and Chapter 6. Before tackling these issues however, the next chapter reviews relevant theoretical models of labour supply, wage determination and wage discrimination which form the crux of the thesis.

CHAPTER 3

3 THEORETICAL MODELS OF LABOUR SUPPLY AND WAGE DETERMINATION

3.1 INTRODUCTION

This chapter reviews a number of theoretical models of labour supply, wage determination and wage discrimination. The purpose is to provide a theoretical basis for the empirical questions answered in this thesis in both a traditional context (i.e. without incorporating health) as well as in a health-augmented scenario. Given that it is not my intention to modify any of the existing theories, the following theoretical review will not attempt to restate the technical details of each model. Rather, attention will focus on the basic meaning and implication of each model.

3.2 OVERVIEW OF LABOUR SUPPLY THEORIES

Theoretical advances in the conceptualization of labour supply have been well documented (Blundell & MaCurdy, 1999; Bowen & Finegan, 1969; Pencavel, 1986). Early theoretical work uncovered an ambiguous wage-labour supply relationship in constrained optimization models (subsequently known to be the result of income and substitution effects of wage variations) (Robbins, 1930). Arguably the earliest study that tested this ambiguity assertion, Douglas (1934) concluded that the elasticity of work hours with respect to wages in the US probably ranged between -0.1 and -0.2. Other important work on LFP include Woytinsky (1940), Bancroft (1958) and Long (1958). Subsequent studies aimed at fleshing out income and substitution effects more carefully, especially from a gendered perspective, were later conducted. For instance, Mincer (1962b) examined the LFP of married women, while Kosters (1966) focused on men's work hours (Heckman & MaCurdy (1981), Heckman et al. (1981) and Killingsworth (1981; 1983) provide a comprehensive review of this literature).

As noted by Pencavel (1986), most of the early work as well as later studies on labour supply draw from Hicks' (1946) model of the determinants of labour supply. These models, which are either based on the individual as an independent unit or view the household in a single-member dimension, or as consisting of a group of individuals seeking to maximize a homogenous well-behaved utility function subject to a family budget constraint, represent the orthodox view (see e.g. Blundell & MaCurdy, 1999; M. R. Killingsworth, 1983; M. R. Killingsworth & Heckman, 1986). Gary Becker was a popular exponent of this conceptualization of household decision-making (Becker, 1965; 1973). These so-called unitary models often assume the existence of a benevolent dictatorial household head who makes decisions on behalf of other household members in the best possible way so as to maximize family welfare.

However, the unitary model has been criticized for some of its apparently unrealistic assumptions and its occasional failure when confronted with data. A key criticism is that the model assumes away key intra-household differences and strategic interaction which have been found to be important in modelling household behaviour at times (Chiappori, 1992; Fortin & Lacroix, 1997). Models that incorporate these important features, otherwise dubbed collective models, often characterize intra-household decision-making in a game-theoretic manner (see e.g. Alderman, Chiappori, Haddad, Hoddinott, & Kanbur, 1995; Fortin & Lacroix, 1997). In this framework, the individual seeks to maximize their own utility within a household set-up, and household resources including time are allocated on the basis of these intra-household power relationships.

Both the unitary and collective labour supply theories draw from the basic static neoclassical model of household behaviour. In this model, an individual/household maximizes a smooth and well-behaved utility function which is a function of goods consumption, leisure/work hours and individual characteristics. Other key properties of the utility function include homogeneity and the satisfaction of other Slutsky conditions like negative semi-definiteness of the Slutsky matrix, while the unitary model includes the assumption of symmetry. According to this framework, the agent's decision problem involves making a trade-off between work and leisure. There is no unemployment as jobs are instantaneously available at the market clearing wage; agents are either employees/labour market participants or non-participants (Mortensen, 1986).

However, the basic static neoclassical model has been criticized as inadequate for explaining labour supply decisions in reality. Important ways through which this model has been enriched include the incorporation of job search and unemployment duration (Mortensen, 1986; Stigler, 1961; Stigler, 1962), labour supply decisions over the entire work life cycle (Blundell & MaCurdy, 1999) and the division of non-market time into both leisure time and time devoted to home production (Becker, 1965).

Though many studies on labour force participation criticize the unitary model as unrealistic, it still remains popular empirically. Perhaps, its continued popularity lies in the fact that much progress has not been made in developing a common framework for collective models that enjoys widespread acceptance among practitioners (Lundberg & Pollak, 1996). And given that the unitary model has been described as a fairly accurate description of the South African situation, it is apparent that it suits our purposes especially given embedded patriarchal forms of family organization in South African families (Ntuli, 2009). This notwithstanding, I am not unaware of empirical evidence which lends credence to the collective theory of intra-family interaction in South Africa (Bertrand, Mullainathan, & Miller, 2003; Duflo, 2003). Thus, given that I make the individual the main object of the analysis while capturing important household-related LFP determinants, I incorporate important aspects of both unitary and collective household decision making into the analysis by modelling LFP as a function of both individual- and household-level characteristics.

3.2.1 Labour supply: theoretical framework

Given the foregoing, this study adopts a standard static labour supply model. The framework adopted here, which closely follows Pencavel (1986) and Blundell and MaCurdy (1999), originates from Hicks (1946).

In this framework, the labour supply function is derived from general consumer demand theory where a fixed endowment of a commodity (a fixed block of time, T) is divided into two parts: one for sale in the market, and another for direct consumption. In this simple characterization, T is to be divided into work hours (h) and leisure time (l), where l incorporates all non-market use of time (including home production). Thus, the agent faces a time constraint, $T = h + l$.

Given perfect certainty, it is assumed that an agent with a "characteristics" vector (X) seeks to maximize a well-behaved utility function (U) defined over her consumption of a composite commodity (C) and work/leisure hours thus:

$$U_i = U(h_i, C_i; X_i, \varepsilon_i) \qquad [3.1]$$

subject to the following constraint:

$$PC_i = Wh_i + Y_i \qquad [3.2]$$

That is, total (labour and non-labour) income is fully spent on the consumption of the composite commodity (C). ε, P, W *and* Y are: unobserved characteristics/tastes/abilities in home production, price of the composite good, the wage rate, and non-labour income respectively. The agent selects $C > 0$ and $h \geq 0$ to maximize (3.1) subject to (3.2). Assuming P to be a numeraire, this yields the following first order conditions:

$U_c(h_i, C_i; X_i, \varepsilon_i) = \mu$ and $U_h(h_i, C_i; X_i, \varepsilon_i) \geq \mu W$, where μ is the marginal utility of income. The marginal rate of substitution between hours of work and goods consumption is the ratio of these marginal utilities.

The agent's reservation wage, W_r is the slope of an indifference curve between commodity consumption and hours of work evaluated at $h = 0$. For each individual, W_r varies for each value of C, and as a result, indirectly upon Y for each X and ε, i.e. $W_r(Y, X, \varepsilon)$. The foregoing implies that an agent's reservation wage is the implicit value of her time when at the margin between participation and non-participation. At this margin, if the market wage (W_m) exceeds the reservation wage, she participates; otherwise, she does not. Consequently, the participation decision is as follows:

$$W_m > W_r \ \Rightarrow participate$$
$$W_m \leq W_r \ \Rightarrow do\ not\ participate \qquad [3.3]$$

These conditions, together with the reservation wage function, $W_r(Y, X, \varepsilon)$ show that relevant determinants of labour force participation include non-labour income as well as observed and unobserved characteristics. In this vein, Sprague (1994) argued that the personal characteristics which determine the reservation wage (and by extension, LFP) include race, gender, marital status, ages and number of children, while the expected market wage offer (W_m) depends on personal and human capital characteristics like age, labour market experience, schooling and unobserved innate ability. Additionally, the general state of the labour market is another key determinant of the market wage offer. This analysis can be easily extended to a multi-member household with minor modifications.

Though the foregoing framework has been extensively used in modelling labour supply/labour force participation as a function of the above-mentioned characteristics, health can be considered a component of the X vector whereby illness is likely to raise one's reservation wage given one's increased taste for leisure. On the other hand, being healthy is likely to increase one's latent productivity and consequently, the opportunity cost of leisure, thereby enhancing willingness to participate in labour market activities. Therefore, the human capital theory hypothesizes a positive relationship between health and LFP (Cai & Kalb, 2006). However, it is possible that participation may occur under perverse conditions due to an unhealthy working environment, thereby resulting in negative health outcomes. These relationships are elaborated in the following theoretical review.

3.2.2 Theoretical review of the health-labour supply relationship

Health is a form of human capital (Becker, 1964; Fuchs, 1972; Mushkin, 1962) while there are many avenues through which both health and LFP (or labour supply in general) can be related. Before detailing these channels of influence, it is important to state that LFP in this context is not only used in the intensive margin (i.e. actual supply of work hours) but also in a broad sense to include the willingness to engage in economic activities of any kind, i.e. broad LFP.

Firstly, being healthy is likely to encourage LFP as it is associated with increased latent and actual productivity. Such higher (latent) productivity associated with better health invariably results in an increase in the opportunity cost of leisure, thereby enhancing willingness to participate in the labour market (Cai & Kalb, 2006). Conversely, sickness increases the demand for leisure and reduces the willingness to participate in the labour market.

Health status can also affect labour supply by changing an individual's preferences between market time and non-market time. Ill-health can affect the manner in which individuals value time away from work thereby changing the relative utility between income and non-market time. More favourable valuation of non-market time by the sick relative to the healthy may stem from the fact that non-market time may be utilized in seeking health care. This channel is independent of the effect of health on participation through changes in productivity earlier mentioned (Chirikos, 1993).

Another important channel through which health may be related to labour market participation is through its effect on life expectancy and the importance of the latter in making labour market decisions. Reduction in life expectancy reduces the time horizon over which an individual makes economic decisions especially as it will imply less need for financial resources to finance years of retirement (van Solinge & Henkens, 2010). Therefore, reduced life expectancy may make early labour market exit more attractive relative to when individuals expect to enjoy longer life spans. This is evidenced by increasing calls to extend the retirement age in the face of longer life spans (van Solinge & Henkens, 2010).

However, it is possible for poor health to result in an increase in labour supply. This is because, conditional on the sick being able to supply actual hours of labour and by reducing productivity, illness decreases earnings and can result in an increase in labour supply. Such increase in labour supply may be necessitated by the need to make up for lost earnings (due to lower productivity) and/or to earn more income given the high cost of seeking health care especially when medical care is associated with direct financial cost (Cai & Kalb, 2006). Thus, though mostly working through channels that predict a positive impact, the direction of the impact of good health on LFP may also be negative and is therefore an empirical question.

The foregoing has demonstrated various channels through which health can influence LFP. Theoretically, LFP can also affect health outcomes in a number of ways. Firstly, employment-related income can be used to procure better nutrition and health care. Secondly, the active nature of participation (especially for those who actually work or are actively seeking employment) can have positive health effects akin to exercise. Moreover, given the value attached by society to gainful employment, non-participation in labour market activities may be associated with psychological trauma due to the shame and stigma associated with non-participation. Indeed, it has been hypothesized that otherwise healthy but non-participating individuals may falsely claim to be sick in order to rationalize their non-participation (Stern, 1989). Such psychological trauma may result in significant health deterioration. Furthermore, lack of activity associated with non-participation may lead to boredom with its attendant adverse health conditions, coupled with social problems like alcohol and drug abuse, depression and suicide (Patterson & Pegg, 1999). On the other hand, participation may also be detrimental to health. Employment in hazardous and stressful environments can result in deteriorating health. An example is work-related health hazards like respiratory disease among miners (Ross & Murray, 2004).

From the foregoing, there seems to be a number of ways through which health is related to LFP or labour supply in general. Health can influence LFP and vice versa. Furthermore, the direction of impact can be positive or negative. The same is true of the effect emanating from LFP. Thus, the relationship is theoretically ambiguous and remains an empirical question. These bi-directional relationships between health and LFP will likely result in endogeneity of health in a labour supply/LFP equation.

3.3 WAGE DETERMINATION: THEORETICAL FRAMEWORK

The discussion in Section 3.2.1 above briefly alluded to factors that determine the market wage (W_m) but did not outline the process of wage determination. Various wage determination theories formulated under the assumptions of perfect and imperfect competition provide insight into the wage determination process. While perfect competition models stress the role of human capital and differences in job characteristics as the reasons for compensating wage differentials, models that incorporate market imperfection stress the role of non-human capital characteristics like barriers to entry, lack of transparency and discrimination as important in determining wage differentials. The following discussion mainly draws from Cahuc and Zylberberg (2004).

Becker's (1964) human capital model provides a foundation for most models of wage determination and accords great importance to the role of human capital, principally education, in driving wage differences. The model fits into the perfect competition framework where it is shown that wages match the supply and demand for labour and equalize marginal productivity in equilibrium. The perfect competition framework crucially depends on the existence of no barriers to entry and exit as well as perfect information about opportunities, worker quality and job characteristics.

Another notable strand in this literature is the hedonic theory of wages. This theory emphasizes job characteristics as important wage determinants.

More difficult jobs will attract higher pay relative to easier ones. An early exposition of this theory was made by Adam Smith who posited that workers with similar capability should be paid differently if their working conditions were different. Rosen (1974) has also made significant contributions to this theory and explains wage differences as resulting from differences in the level of difficulty associated with different jobs. Within the hedonic theory of wage differentials under perfect competition, two scenarios can be examined: a situation with jobs of equal difficulty; and another, with jobs of different difficulty levels (Cahuc & Zylberberg, 2004). In the former, the (omniscient) planner's problem is to choose levels of worker disutility of work to maximize the sum of individual utilities. The solution to this problem is the equalization of work disutility with worker's marginal productivity.

But assuming different levels of difficulty across jobs, where jobs with higher difficulty/risks are more productive, the worker's problem is to choose an effort level to maximize utility (a function of income, effort level and work aversion) subject to participating in the labour market. The worker only participates if the marginal returns to effort equals the job's disutility. The wage received by a worker in equilibrium is therefore equal to the productivity associated with the job, where productivity is a function of effort exerted. This scenario yields multiple equilibria, where individuals with high aversion for work choose jobs requiring little effort, thus earning low wages in equilibrium while the converse obtains for those with low aversion for work. An implication of this theory is that public policy aimed at mitigating the difficulty associated with jobs in a perfectly competitive setting will lead to welfare loss.

However, the actual process of wage determination rarely takes place in the controlled world of perfect competition. Principal reasons for deviations from perfect competition are information asymmetry and entry barriers. Market imperfection results in wage differentials that are not solely driven by differences in marginal productivity (Cahuc & Zylberberg, 2004). Factors that impinge on workers' ability to freely move so as to earn their marginal productivity include geographical and occupational factors. Transport costs for instance, may prevent a worker from taking advantage of work opportunities elsewhere other than their immediate environment even when the local wage does not equalize marginal productivity. On the other hand, costs of entry sometimes prevent firms from entering a market dominated by a monopsony or oligopsony (Cahuc & Zylberberg, 2004).

The foregoing review has identified education, information asymmetry, job difficulty and entry barriers (due to geographical and occupational barriers for instance) as some of the factors that can theoretically determine wages. Additionally (and particularly relevant to this thesis), health can be considered an important determinant of productivity/wages.

The idea that health is an important determinant of productivity stems from the recognition that it is a form of human capital since it affects the ability to produce output (Weil, 2007). And it is an established fact that better human capital boosts productivity and hence wages, ceteris paribus. However, the point that health has a human capital element does not negate the fact that some important aspects of health are the result of genetic, cultural and social factors which are generally pre-determined. It only implies that some aspects of human capacities contribute to increased output and can be influenced by human choices (Savedoff & Schultz, 2000; Weil, 2007).

Research into human capital formation was boosted by Becker's (1964) observation that growth in physical capital and labour contributed relatively little to income growth in most countries. The earlier growth accounting framework of neoclassical economic theory identified capital, labour and productivity as the key drivers of economic growth, where productivity was believed to be driven by technological progress, considered exogenous in the neoclassical framework (Solow, 1956). Subsequent advances modelled technological progress as endogenously determined by human capital investments, chiefly in education, resulting in skilled labour. Becker asserted that human capital investment raises productivity, thus accelerating economic growth.

Though much emphasis on the role of human capital has focused on education, there has been increasing attention on health as a veritable component of human capital. As Suhrcke et al. (2005) observed, it is intuitively obvious that health would impact on labour productivity. Jack and Lewis (2009) maintained that the most obvious reason why healthier people are more likely to be richer than the sick is their greater capacity to work harder, longer and more consistently than the latter. Deaton (2002) also agreed that good health enhances one's ability to work, thereby enhancing productivity and wage. Other channels of influence include the fact that poor health induces early retirement (Bazzoli, 1985; Bound, Schoenbaum, Stinebrickner, & Waidmann, 1999; McGarry, 2004), presupposing subsequent loss of wage income. There is also a long-term and inter-generational link between poor health and future productivity. Early parental death due to illness results in orphans whose educational prospects, future productivity and earnings are severely hampered (Case et al., 2004; Jack & Lewis, 2009). Similarly, mothers' smoking habits during pregnancy have been found to adversely affect their children's health, educational attainment and subsequent productivity, while stunting (partly due to say, childhood diseases) has been associated with higher future unemployment (and therefore lower lifetime wage income) (Deaton, 2002). Furthermore, in utero nutritional supplements can affect children's future human capital acquisition; for instance, reduction in iodine deficiency disorders through intensive iodine supplementation of mothers resulted in additional 0.35-0.56 years of schooling for the benefitting children in Tanzania (Field, Robles, & Torero, 2009).

Further justification of the importance of health in driving productivity and wages stems from efficiency wage theories which highlight the role of nutrition in engendering improved health. For instance, Leibenstein (1957) hypothesized that better nourished workers are more productive than their poorly nourished counterparts while the gradient between nutrition and productivity is steeper at very low levels of nutrition. The implication of such efficiency wage theories is that employers have an incentive to pay healthy workers above the minimum wage while excluding those in poor health from the labour market given their high cost of employment (Strauss & Thomas, 1998).

A popular and clear formulation of the human capital nature of health can be attributed to Grossman (1972) in a modification of Becker's human capital model (Becker, 1964). The Grossman model shows that current health status is a function of previous health stock, material and time inputs into health production, exogenous productivity shifters (such as the rate of depreciation of health capital which likely varies with age[5]), person-specific productivity shifters (e.g. health investments) and random productivity shock. It clearly identifies the relationships between health, work time and wages. According to the model, individuals produce as well as demand health. In terms of health production, individuals are born with a stock of health which deteriorates over time through sickness for instance but replenished via investments in health care. Also, people demand health both as a consumption and capital good. As a consumption good, health directly enters the utility function as individuals enjoy being healthy. As a capital good, it is used in the production of healthy time which can be spent in both market and non-market activities, where healthy time increases both time spent in productive work as well as productivity, ceteris paribus. Therefore, the marginal utility of having an extra unit of health stock has both a consumption and capital component, the sum of which must equal the marginal cost in equilibrium (Zweifel, Breyer, & Kifmann, 2009).

Though health potentially influences productivity/wages, wage no doubt affects health. That wage is a predictor of health stems from the health-income literature. Higher wages facilitate the purchase of essential nutrients and services important for healthy life as well as the ability to control one's circumstances (see e.g. Marmot, 2002). Also, a dominant income effect of

[5] In a more complicated version of the model, Grossman treated the rate of depreciation as endogenous, but noted that its exogenous treatment greatly simplified the model, while both versions yielded similar conclusions.

wage increases over the substitution effect results in the well-known backward-bending labour supply curve in neoclassical economic theory, leading to reduced labour supply and enjoyment of more leisure - when leisure is a normal good - (with its attendant health benefits) than would have been possible at low wages (Pencavel, 1986). But given this chapter's focus on the gradient running from health to wages only, I will not delve into this literature. A concise theoretical formulation of this endogenous relationship is available in Savedoff and Schultz (2000).

3.4 THEORIES OF DISCRIMINATION

Though the above theories of wage determination have explained wage differentials as emanating from differences in skill/preferences of workers, some important extensions have established the possibility of wage differentials between otherwise productively identical workers (Altonji & Blank, 1999). This is a classic case of labour market discrimination. Discrimination may however be very difficult to detect. This is because discrimination can influence human capital investment decisions pre- and post-labour market entry (e.g. when women acquire less education or select themselves out of certain disciplines due to discrimination) -see e.g. Coate and Loury (1993). In the following section, I discuss factors that can theoretically result in wage discrimination.

Defining labour market discrimination as a situation whereby providers of labour services who possess equal levels of physical or material productivity are unequally treated in a manner related to an observable characteristic like gender, ethnicity or race, Altonji and Blank (1999) identified two broad classes of economic models of discrimination: *competitive* models (which focus on agents' individual actions), and *collective* models (which stress a group's collective actions against another). Competitive models highlight two broad kinds of discrimination. The first is *prejudice*, defined as a taste by some or all members of a majority group against having interaction with those of a minority group[6], while the other is *statistical discrimination* by employers due to imperfect information regarding the skills or behaviour of minority group members. On the other hand, collective models pay more attention to the consequences of *collective action* of one group against another. Such models are often informal and emphasize the use of the legal system or the threat of violence as an enforcement mechanism (Altonji & Blank, 1999). The focus in this review is on competitive models as they are more popular among mainstream labour economists and also adequately illustrate the point intended in this thesis, i.e. that disability/impairment can be a marker of discrimination.

3.4.1 Prejudice/taste-based discrimination

Becker (1971), who modelled prejudice as a taste for discrimination, defined employer discrimination as a situation in which some employers are prejudiced against a minority group. In this model, employers seek to maximize a utility function which is the sum of profits and the monetary value of utility derived from employing members of particular groups. Prejudiced employers can only hire workers from the minority group if the wage paid to the dominant group is at least equal to the prejudice-adjusted wage rate offered the former (where the prejudice-adjusted wage rate exceeds the nominal wage rate payable to the minority group in the absence of discrimination). The implication is that only the least prejudiced firms employ workers from the minority group, thus resulting in a segregated labour market between majority and minority groups. Furthermore, the model implies that non-discriminating employers would earn higher profits than their discriminating counterparts since non-discriminators would pay less for hiring minority workers.

The model predicts the elimination of the wage gap in the long run if there is free entry and/or constant returns to scale. This is because in the long run, the number of non-discriminating employers would have grown to the point

[6] Minority and majority groups are used throughout this discussion to refer to the group being discriminated against and those being favoured respectively, and do not necessarily connote numerical strength.

where minority workers need not work for prejudiced employers. However, this is in stark contrast to the persistence of, say, gender- and disability-related wage gaps found in empirical studies. The inability of the theory to concur with empirical evidence suggests a number of reasons. Either there is no more discrimination (implying that the observed gaps are caused by productivity-related factors) or all potential employers are discriminators, or that employer discrimination is actually not the main form of labour market discrimination, or still that other factors such as collective action or search frictions impinge on the expansion of non-discriminating firms. This model can be extended to account for employee discrimination, consumer discrimination as well as taste-based discrimination in the presence of costly search (Altonji & Blank, 1999).

3.4.2 Statistical discrimination theory

Most of the literature on labour market discrimination has been devoted to explaining the consequences of race- or sex-based statistical discrimination. The main idea of statistical discrimination models is that firms possess limited information with regard to the applicants' skills and turnover propensity. Therefore, these firms may resort to using easily observable characteristics like disability status or ethnic origin to statistically discriminate among employees if it is believed that these characteristics are correlated with productivity. Major contributors to this literature include Phelps (1972) and Arrow (1973).

Two main strands of the statistical discrimination theory have emerged. One is concerned with ascertaining how employers' prior beliefs/stereotypes regarding different group members' productivity affect their hiring and remuneration decisions. Arrow (1973) and Coate and Loury (1993) have made significant contributions to this aspect of the literature. The other strand is concerned with issues regarding the consequences of group differences in the level of the precision of information that firms have with regard to individual productivity. Prominent contributors to this literature include Aigner and Cain (1977), Lundberg (1991) and Oettinger (1996).

On the role of employers' stereotypes, Coate and Loury's model shows that differences in firms' prior beliefs about the skills of different groups of workers can result in equilibria whereby groups that possess the same level of innate ability end up with different levels of skills. In this model, all workers have the same level of basic skills and employers observe workers' group membership as well as a noisy signal ratio (whose distribution is a function of the worker's qualification status). In deciding whether to assign the worker to a skilled or unskilled job, the firm forms a posterior probability that the worker is qualified based on the observed signal as well as a prior belief that a member of that particular group is qualified, assigning all workers above a given critical value of the posterior probability to the skilled job. On the other hand, workers only decide to become trained if the value of the change in the probability of being assigned the highly skilled job exceeds the cost of training. Overall, the model predicts that even if both workers' groups have identical skills and the same distribution of training costs and firms update their priors in a sensible way, stereotypes that are initially negative may become self-confirming, i.e. prior beliefs against a particular worker group will result in a lower wage for members of that group.

As mentioned earlier, another strand of the statistical discrimination theory deals with issues regarding the consequences of group differences in the level of the precision of information that firms have with regard to individual productivity. Lundberg's (1991) model (which is prominent in this literature and builds on earlier models- see e.g. (Altonji & Blank, 1999)) starts by assuming that the accuracy of the information firms have about the productivity of individuals differs across groups. The model shows that this can result in an equilibrium in which firms statistically discriminate on the basis of group membership and there are ex post group differences in productivity, though average innate ability is the same across groups (Altonji & Blank, 1999). In a situation where firms are allowed to discriminate, a firm will use two separate wage equations for the two worker groups given the same training cost parameter and mean innate ability across groups (thereby paying different wage rates across groups). If firms are not allowed to discriminate given group-specific productivity parameter, differences in wages and human capital investment will be eliminated. Also, preventing firms from estimating different productivity equations for the two groups will reduce the accuracy of firms' estimates of productivity (Lundberg, 1991).

3.5 CONCLUSION

This chapter has reviewed models which provide a theoretical basis for studying the determination of LFP, wages and wage discrimination. Key points that emerge from this discourse are that unearned income and human capital characteristics like health, education and job characteristics are important determinants of labour supply. Also, human capital characteristics and tastes result in wage differentials. On the other hand, even equally productive individuals may be paid differently due to factors unrelated to their productivity. Such factors might include one's group membership. The following three chapters ascertain whether differences in health status help explain these labour market phenomena empirically in South Africa.

CHAPTER 4

4 THE IMPACT OF HEALTH ON LABOUR FORCE PARTICIPATION IN SOUTH AFRICA

4.1 INTRODUCTION

It was pointed out in Chapter 1 that the non-incorporation of health into most labour supply models is a gap in the South African labour supply literature. In Chapter 2, I showed that ill-health was one of the main reasons why South Africans were economically inactive in both 2008 and 2012 (Table 2.4). Consequently, this chapter examines the impact of health on LFP in South Africa. This is essential because if ill health results in reduced LFP by working age economic agents, then it imposes a cost on the economy beyond the cost of treatment. Additionally, the magnitude of such impact is an important input into the conduct of cost-benefit analysis of interventions required to boost population health from an economic point of view (Cai & Kalb, 2006).

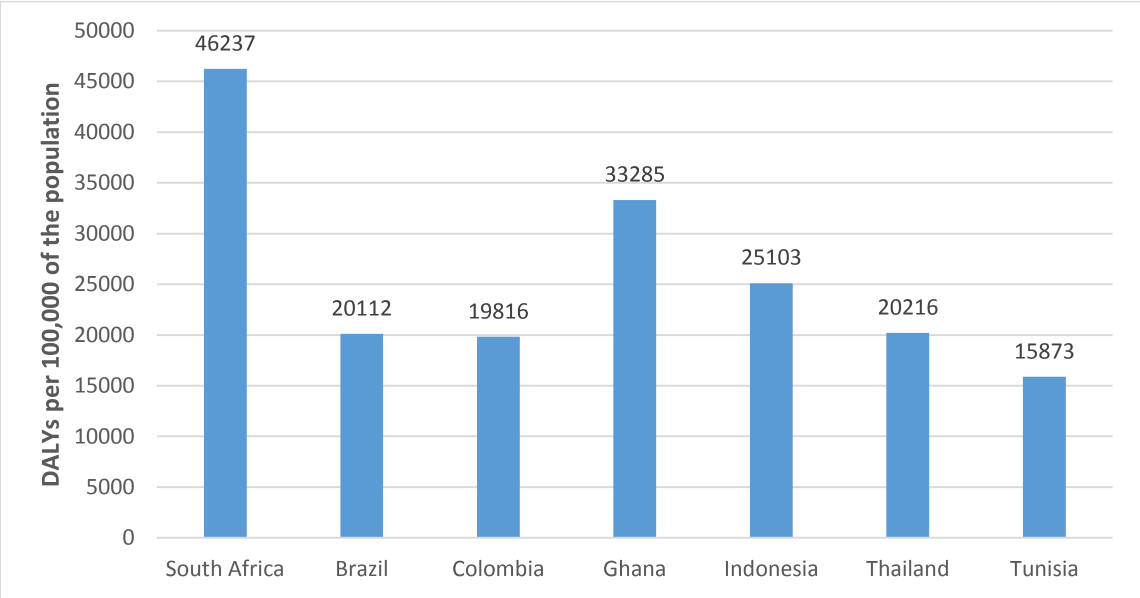
Though a number of studies have examined the trends and determinants of labour supply in South Africa, the possible impact of health is largely lacking (Banerjee et al., 2008; Bhorat, 2007; Ntuli & Wittenberg, 2013; Wittenberg, 1999). Available data show that prime-age LFP has been declining in recent years in South Africa; for instance, though strict LFP rate increased from 51.4% to 59.4% between 1995 and 2001 in South Africa, it declined to 57.2% in 2005 (Banerjee et al., 2008), with a further decline to 54.3% in the final quarter of 2011 (Statistics South Africa, 2012a). Given the important role that gainful labour market participation plays in economic growth as well as the role of the labour market in driving income inequality especially in South Africa (Leibbrandt & Woolard, 2001), it is imperative to examine at least most of the factors which might be responsible for a decline in participation.

Also, though not the focus of this chapter, demand conditions are very important in understanding labour market dynamics in an economy. In South Africa, it has been observed that together with increased labour supply, especially in the post-apartheid period, a virtually stagnant labour demand has greatly contributed to the high unemployment rates characterizing the country (Banerjee et al., 2008). Given that it is not my intention to rehash the earlier discussion about labour demand in South Africa, suffice it to say that labour demand factors such as growth in capital-intensive, as opposed to labour intensive sectors, the general low labour absorption in the formal economy, as well as poor growth of the informal economy, are also essential in determining labour supply outcomes. This is because poor demand conditions may discourage job search efforts, thereby leading to declining labour supply.

Moreover, South Africa is beset with substantial disease burden as a result of the high incidence of HIV/AIDS and rising morbidity from non-communicable diseases (NCDs) (Bradshaw et al., 2000). Rising mortality from NCDs especially is not surprising given some lifestyle choices among South Africans. For instance, 35% of adult men smoked tobacco while 48% (63%) of adult men (women) were categorized as inactive in 2003; furthermore, most age groups from 15 years upwards reported an increase in tuberculosis incidence while the prevalence of hypertension also increased between 1998 and 2003 (Department of Health, Medical Research Council, & OrcMacro, 2007). Also as Figure 4.1 shows, South Africa had a high disease burden in 2004 even relative to other select developing countries. And compared to 2000 disability-adjusted life year (DALY) figures (approximately 16 million), 2004 figures (approximately 22 million[7]) represent substantial increase in DALYs even in the face of (at least slightly) declining HIV/AIDS incidence (AVERT, undated), suggesting increased morbidity from non-HIV/AIDS-related causes.

[7] Calculated from the South African DALY in Figure 4.1 using a total population of 48 million.

Figure 4.1: Absolute disease burden across select developing countries, 2004



Source: ECONEX calculations from WHO (2009)

In a nutshell, South Africa has experienced high (and arguably increasing) disease burden in about the same period characterized by declining LFP in the country. Therefore, it is important to investigate if health influences LFP in the South African context.

4.2 RESEARCH AIM AND OBJECTIVES

The aim of this chapter is to ascertain the impact of health on LFP. Specifically, this chapter ascertains:

i) Whether self-assessed health (SAH) exerts a significant contemporaneous effect on LFP in South Africa.

ii) Whether the relationship between SAH and LFP extends beyond the contemporaneous period in South Africa.

4.3 HEALTH STATUS AND LABOUR SUPPLY: CONCEPTUALIZATION AND METHODOLOGICAL ISSUES

Researchers are often faced with the problem of deciding what health measure to use in a labour supply model. Apart from observed health status being possibly endogenous due to perceived simultaneity between both health and labour supply, most of the measures used to represent health suffer from measurement problems given their imperfection in accurately capturing work capacity (see e.g. Lambrinos, 1981). This is because health is like ability, in the sense that inasmuch as everybody has an idea of what it means, it remains difficult to measure (Currie & Madrian, 1999). This failure to properly measure health potentially leads to a bias akin to what Griliches (1977) termed ability bias in standard human capital models. Thus, the definition of health has proved to be a vexing issue. Perhaps the most popular definition is that of the WHO: a state of complete physical, mental and social wellbeing, and not merely the absence of disease or infirmity (WHO, 1946, p. 100).

However, Saracci (1997) has criticized this definition as being more idealistic than practical especially with regard to informing researchers of how to actually measure health. He maintained that the above definition is related more to happiness than health, where clear distinctions exist between both concepts (for instance, one may engage in a health-enhancing activity which results in a decline in happiness, such as quitting smoking). Therefore, he proposed an alternative and more practical definition which attempts to separate both concepts, where health is defined as a condition of wellbeing, free of disease and infirmity and a basic and universal human right.

The foregoing demonstrates the difficulty inherent in measuring health. The incomplete understanding of the full ramifications of health has led to the use of health proxies such as life expectancy and infant mortality ratio (Pritchett & Summers, 1996), illness severity and body mass index (BMI, measured in kg/m^2)[8] (Rivera & Currais, 2005), SAH measures (Cai, 2010),

---

[8] $\frac{weight\ (kg)}{height^2(m)}$

disability measures/functional limitations with activities of daily living (ADL) (Stern, 1989), and disability-adjusted life years (DALYs) (Murray, 1994) among others in exploring the health-labour supply relationship. As noted by Cameron and Trivedi (2005), such measurement errors if not properly accounted for, may attenuate estimates of the impact of the covariate of interest (in this case, health) on an outcome (say LFP), thus rendering it inconsistent. The foregoing measurement issues have led to vigorous debate on which health measures are a better reflection of true unobserved health status.

It has been argued that more comprehensive health measures (e.g. whether health limits work, and self-reports of overall health status) increase the explanatory power of health in labour supply models compared to the use of limitation with a specific activity of daily living (Manning, Newhouse, & Ware Jr, 1982). But as shown subsequently, some studies prescribe the use of relatively objective health measures (e.g. life expectancy, BMI and limitations with ADL) over more subjective measures like SAH (see e.g. Kreider, 1999). This is because the measurement error associated with self-reported health might not be random, where non-labour force participants may be more likely than participants to cite illness as the reason for their non-participation given the social stigma associated with non-participation and the fact that receipt of certain public transfers is dependent on health status (Boskin, 1977; Currie & Madrian, 1999; Parsons, 1980). This potentially leads to a bias referred to as rationalization endogeneity, a bias likely to lead to an over-estimation of the impact of health and under-estimation of the effect of financial variables on labour supply (Bound, 1991; Bound, Schoenbaum, & Waidmann, 1995; Cai, 2010; Cai & Kalb, 2006). Furthermore, self-reported health measures might be influenced by whether or not the individual in question has sought treatment (Dow, Gertler, Schoeni, Strauss, & Thomas, 1997), where utilization of medical services generally increases with income despite the fact that the rich (who often have a high LFP probability) are generally healthier than the poor (Currie, 1995; Strauss & Thomas, 1998). This scenario can result in a spurious positive relationship between reported ill-health and labour supply.

The potential for subjective/self-reported health measures to suffer from rationalization endogeneity has led to some authors preferring the use of objective health measures (see e.g. Bartel & Taubman, 1986; Johnson & Lambrinos, 1985; Kreider, 1999; Mitchell & Burkhauser, 1990). However, to the extent that these objective health measures are imperfect measures of true health status, they are likely to under-estimate the impact of health on labour supply. For instance, height, which reflects quality of childhood nutrition and the absence of debilitating disease conditions especially during one's formative years (Fogel, 1994; Savedoff & Schultz, 2000) may be a poor reflection of work capacity (e.g. computer programming ability) even though its associated measurement error is likely to be random in a population (Currie & Madrian, 1999; Strauss & Thomas, 1998). Also, subsequent mortality has been criticized as inadequate in capturing true health status even with most available longitudinal data (Bazzoli, 1985). Therefore, objective health measures suffer from errors-in-variables bias just as subjective ones. It should be noted that objectivity/subjectivity may be understood in relative terms as even self-reported health conditions referring to some specific condition or activity (e.g. walking, bathing, and seeing) may be considered objective relative to a self-reported assessment of overall health status even as the former may not be based on physician diagnosis.

Given the foregoing, Bound (1991) suggests that subjective health measures might be associated with less bias than objective ones since they would likely be affected by two opposite sources of bias (rationalization endogeneity and error-in-variables) which might cancel out, while objective health measures are only likely to be biased downwards (due to the attenuation bias associated with errors-in-variables bias). As noted by Currie and Madrian (1999), this is consistent with the observation of smaller health effects in labour market studies that use objective health measures relative to those that employ subjective ones. Even more important at least for the purposes of this study is the fact that a wide range of studies covering different age groups, geographical entities (including South Africa) and patient groups have affirmed that global SAH measures predict subsequent mortality fairly well (Ardington & Gasealahwe, 2014; DeSalvo, Bloser, Reynolds, He, & Muntner, 2006; Frankenberg & Jones, 2004; Larsson, Hemmingsson, Allebeck, & Lundberg, 2002; Nybo et al., 2003; Wong, Wong, & Caplan, 2007). Such a strong relationship between SAH and mortality apparently results from the largely logical context informing SAH responses as well as its relatively comprehensive nature. Indeed, SAH has been demonstrated to emanate from respondents' rational thought processes, while encompassing various dimensions of their health (including cultural and biological), as well as bodily sensations which may not be easily detected via clinical tests (Jylhä, 2009).

But in proxying true health status with subjective health measures in labour supply models, some researchers have advocated instrumenting the subjective health measure with relatively objective health indicators as a means of overcoming the potential inconsistency of their estimates. Such objective instruments include subsequent mortality (Anderson & Burkhauser, 1984; 1985; Parsons, 1982) and specific health conditions and symptoms like general weakness, blindness and breathing problems (Campolieti, 2002; Dwyer & Mitchell, 1999; Stern, 1989; Haveman, Wolfe, & Huang, 1989). However, Bound (1991) maintained that this may not be enough to ensure consistency. If the correlation between SAH and the error term in the LFP equation (i.e. rationalization endogeneity) is (at least partly) due to the correlation between SAH measurement error and, say, financial variables in the LFP equation, then even instrumenting SAH with objective health conditions a la Stern (1989) will yield inconsistent estimates of the effects of both SAH and the financial variables on LFP. This is the case when, say, the unemployed over-report illness as the cause of their unemployment status in expectation of unemployment-related financial rewards predicated on illness. But as noted by Au et al. (2005), if the correlation between SAH and the error term in the labour supply equation is only due to the correlation between the measurement error in SAH and unobserved determinants of labour supply (e.g. taste for leisure), then instrumenting SAH with objective health measures yields a consistent estimate of the impact of health on labour supply, even though the coefficients of financial variables will still remain inconsistent.

The foregoing discussion shows that no health proxy strictly dominates the other a priori in capturing the true effect of health on labour supply. That said, it is likely that composite health measures like SAH might be more effective in uncovering the health-labour supply relationship compared to individual illnesses/symptoms. Given the foregoing, this chapter used SAH to proxy true health status.

4.4 HEALTH STATUS AND LABOUR SUPPLY: AN EMPIRICAL REVIEW

Many studies have investigated the relationship between health and labour supply. The main labour supply outcomes often researched in the literature include LFP (extensive margin), employment probability and number of hours worked. These outcomes are surely not only determined by supply factors; demand factors such as whether job opportunities exist in the first place as well as the general economic condition play an important role in their determination and can be particularly important in a country like South Africa characterized by high unemployment rates. In this review, I focus on the labour supply characteristics of these outcomes in line with the way they are conceptualized in the reviewed literature.

Stern (1989) examined the simultaneous impact of health (SAH and health limitations) on LFP and vice versa in the US using the 1978 Survey of Disability and Work and the 1979 cohort of the Health Interview Survey. The health and labour participation equations were identified by excluding marital status and its interaction with sex from the former, and various health conditions (like blindness, weakness and walking problems) from the latter. Results showed that self-reported disability status significantly explained LFP while there was evidence of weak endogeneity of disability on LFP. Also, there was no evidence of systematic over-reporting of disability by non-labour force participants.

However, a drawback of the study is the difficulty in justifying the exclusion of marital status from the health equation as well as the inadequate test of exogeneity of SAH. While showing that better SAH increased the probability of LFP for all sex-age groups, similar studies conducted in Australia by Cai and Kalb (2006) and Cai (2010) rejected the hypothesis of SAH exogeneity. To achieve identification, Cai and Kalb excluded birth place, presence of young children and interactions between the latter and marital status from the SAH equation, while occupation, work status, smoking and drinking status, long term illness, lack of physical activity, and physical functioning were excluded from the LFP equation. For Cai (2010), only the presence of young children and its interaction with marital status were excluded from the SAH equation, with only the presence of a long term illness and physical functioning excluded from the LFP equation, perhaps realizing the difficulty in justifying the exclusion of smoking and drinking habits from labour supply equations as noted in Heckman et al. (2006). Also, Wilkinson (1995) and Haan and Myck (2009) suggested the existence of a significant relationship between having young children and parental health, thus rendering the use of the former as a means of identification suspect.

Other studies that have used health conditions and/or some measure of physical functioning as health instruments include Dwyer and Mitchell (1999) and Campolieti (2002). The case for using them as instruments for SAH stems from human capital theory where health is seen as a form of human capital. Thus, it is taken that adverse health conditions/limitations in physical functioning affect participation in the labour market only through their effects on the underlying health capital (here proxied by SAH).

In estimating the effect of health on the work effort of single mothers in the US, Wolfe and Hill (1995) found that a single mother's limitation with respect to ADL, whether she reported poor/fair health compared to other health assessments, and whether she had a disabled child, significantly reduced her probability of being employed. However, the exogenous treatment of these health variables (without formally testing the suitability of such an assumption) makes the interpretation of these coefficients as causal effects suspect.

Using data from the Health and Retirement Study (HRS), Kreider's (1999) simultaneous model of work participation, disability and income flows opposed the use of self-reported health status as a reliable measure of true health in labour supply models. The health variable was a three-category ordinal variable representing various levels of health-induced work limitation. Work limitation of only workers was used to predict that of non-workers given the author's belief that the former had no incentive to over-report disability-related work limitation. The results suggested that male (female) non-workers over-reported work limitation by 2.8 (6.2) standard deviations, while over- reporting was most prevalent among non-working high school drop-outs, non-whites and former blue collar workers. Similarly, Kreider and Pepper (2003) found that models estimated under the assumption of accurate disability reports of both workers and non-workers led to biased inferences, with non-workers having a tendency to over-report disabilities. But a concern regarding Kreider's (1999) methodology is that even if non-labour force participants are more likely than participants to use health as an excuse for non-participation, it is not clear that labour force participants'/workers' work limitation responses are accurate. Currie and Madrian (1999) have observed that the latter might under-report their health limitations. Also, O'Donnell (1998) has noted that Kreider's estimates lack efficiency since information on work limitation was only obtained from workers.

In contrast to Kreider (1999), Benitez-Silva et al. (2004) found, using the same data (i.e. HRS) that disability applicants did not over-state their disability status on average. Even though Au et al. (2005) found some evidence of rationalization endogeneity when using SAH to determine employment in Canada using the Canadian National Population Health Survey, employment effects of disability were similar to what obtained when more objective health measures were used. Also, using the National Population Health Survey in Canada, Campolieti (2002) found a large negative effect of disability on the LFP of older Canadian males but suggested that using self-reported health measures will under-estimate the impact of disability on labour force decisions.

Though O'Donnell (1998) concurred with Kreider's finding that non-workers' disability responses were unreliable, he suggested that such a conclusion should only be made after conducting formal tests. In his work on the effect of disability on employment using the British Office of Population Censuses and Surveys which comprised only disabled respondents, O'Donnell made the case for recognition that disability potentially leads to work incapacity. In other words, models of disability and employment should incorporate the fact that employment is determined by both a capacity and desire for work, and not only the latter. He contended that the employment status of some disabled individuals was not the outcome of a comparison of utility in two states: employment and non-employment, but rather that only one state was possible (i.e. non-employment due to work incapacity). This is supported by Nagi (1969) who observed that 55% of US disability insurance applicants were not fit for any kind of work, as well as Bound (1989) who concluded that many disability insurance applicants would not be working even in the absence of the disability insurance programme because of the extent of their disability.

Using the 1997 and 2003 waves of the UK Labour Force Survey data, Jones et al. (2006) explored the relative effects of physical and mental disability on employment. They found that disabled respondents (work-limited and non-work-limited) with physical disabilities were more likely to be employed than their counterparts with mental disabilities. However, their inability to account for the likely endogeneity of health in employment models due to lack of instruments makes such a conclusion tenuous.

Gomez and Nicolas (2006) employed difference in difference and matching techniques in studying the simultaneous causal relationship between health (i.e. SAH) shocks and employment in the Spanish labour market using the European Community Household Panel. To avoid possible rationalization endogeneity, only individuals who reported health deterioration in the second and third years of a three-year cycle but remained employed in the first and second years were counted as truly having experienced a health shock. Likewise, only those who were employed in the first year, unemployed in the last two years but reported being in good health in the first two years were regarded as genuinely having an adverse employment shock. They found that individuals who suffered a health shock were around 5% less likely to remain in employment and 3.5% more likely to remain inactive than those who did not suffer any health shock.

Haan and Myck (2009) examined the dynamic (sequential) simultaneous relationship between health (SAH) shocks and non-employment risk of adult males aged 30-59 years using the 1996-2007 waves of the German Socio-economic Panel. Similar to Bartel and Taubman (1986) and Haveman et al. (1994), lagged health status (non-employment) was used to predict current non-employment risk (health status) in order to avoid endogeneity. Results showed that lagged SAH (non-employment risk) exerted a significant impact on current non-employment (SAH) whether or not unobserved heterogeneity was controlled for.

Some studies have also tracked the relationship between health status and future labour market outcomes. Using the Panel Study of Income Dynamics (PSID), Smith (2009) found that having self-reported one's health to be excellent or very good up to age 16 had a positive and statistically significant effect on future hours of work relative to worse health outcomes. Also in a study that explored own/sibling's labour market implications of health outcomes using the relatively income poor sub-sample of the PSID, Choi (2007) found that an unhealthy young adult was less likely to work as an adult, while females with unhealthy siblings in young adulthood were more likely to work as adults relative to other females, suggesting long term effects of illness and perhaps intra-family income transfer requirements.

To the author's best knowledge, there is not much evidence on the relationship between health and labour supply in developing countries. In Taiwan, Mete and Schultz (2002) found a negative association between poor health and LFP among the elderly. Bridges and Lawson (2008) also found that poor health was associated with a decline in the probability of being in the formal labour market in Uganda, where this relationship was stronger among women. Moreover, investigating the relationship between health status and LFP in Sub-Saharan Africa using a dynamic panel data model with 46 countries, Novignon, Novignon and Arthur (2015) found a positive and significant relationship between population health and LFP in the general and female populations.

Relatively few studies have investigated the health-labour supply relationship in South Africa. They include Arndt and Lewis (2001), Booysen et al. (2002), Young (2005) and Levinsohn et al. (2013). While Arndt and Lewis found no significant impact of HIV/AIDS on employment in South Africa, Young found a significant long run relationship. Booysen et al. also found some relationship between the disease and labour supply in the Free State Province. Apart from not utilizing nationally representative survey data, these studies largely ignored the causal dimension of labour supply determination in South Africa. These concerns were addressed by Levinsohn et al. who found, using propensity scores, that being HIV-positive was associated with a 6-7 percentage point increase in unemployment probability. A common feature of these studies is that none evaluated the impact of a composite/global health indicator like SAH or limitations with activities of daily living on these labour market outcomes as each focused on a single health measure, mainly HIV/AIDS. Consequently, this study used IV techniques to ascertain the impact of a global health measure (i.e. SAH) on LFP as opposed to the non-causal analysis done by most of these studies.

4.5 ECONOMETRIC ANALYSIS

4.5.1 Modelling strategies and theoretical model specification

Firstly, I estimated “traditional” LFP equations (i.e. without including health as a regressor). Subsequently, the bivariate probit model and instrumental variable linear probability model (IV-LPM) were estimated (given the apparent endogeneity of SAH as well as the binary nature of both observed LFP and SAH). Thereafter, the statistical properties of the models were ascertained. Though a case was made for the suitability of the instruments including the fact that similar instruments have been used in previous empirical work, I am not unaware of the difficulty in obtaining convincing instruments in social science research. Therefore, a number of models were estimated to ascertain the existence of a relationship, however non-causal, between health and LFP in both contemporaneous and temporal settings. Finally, given the apparent gender bias in labour market outcomes in South Africa as shown in Chapter 2, the analysis was also disaggregated by gender.

Though the bivariate probit model is an important model with which to recover the effect of a dichotomous variable on a dichotomous outcome as in this study, its dependence on normality of the joint distribution of the error terms necessitates incorporating IV-LPM into the analysis as the latter is not reliant on the joint normality assumption for consistency. However, two well-known shortcomings of the LPM are heteroscedasticity and probability predictions outside the unit interval. In this application, the former is not a major problem given that all estimates were corrected for heteroscedasticity. Also, most of the predicted probabilities in the model lie within the unit interval; indeed, there was no negative predicted probability while only 2.4% of predictions exceeded 1. Among those that exceeded 1, the average value was only 1.02 while the largest value was 1.07, thus suggesting that these predictions are not likely to exert significant influence on the estimates. IV-LPM has also been shown to be quite robust and simple in implementation and interpretation (Angrist & Pischke, 2009).

In the first stage, I assume that latent health status (h_i^*) is a linear function of a vector of exogenous variables, X (e.g. socio-economic status) and excluded instruments (z_i) necessary to identify the impact of health on LFP, and a random error term ($\varepsilon_{i,H}$). In the second stage, unobserved latent LFP (l_i^*) is specified as a function of a vector of exogenous variables (X), latent health status (h_i^*) and a random error term, $\varepsilon_{i,L}$. Though z_i is an instrument vector, it is represented as a scalar in equation (4.1) below for notational simplicity. Importantly, there are possibly some unobserved joint determinants of both latent LFP and health status such as genetic heterogeneity. These relationships can be modelled in both a bivariate probit framework, assuming joint normality of the error terms, $\varepsilon_{i,H}$ and $\varepsilon_{i,L}$ (see Section 4.6.3 of Angrist and Pischke (2009) and Section 15.7.3 of Wooldridge (2002)) or in the two stage least squares framework using the IV-LPM framework. The relationships are specified as follows:

$$h_i^* = X^/\beta^* + \gamma^* z_i + \varepsilon_{i,H} \qquad [4.1]$$

$$l_i^* = X^/\alpha^* + \delta^* h_i^* + \varepsilon_{i,L} \qquad [4.2]$$

Model assumptions in the bivariate probit framework are as follows: $E(\varepsilon_{iL}) = E(\varepsilon_{i,H}) = 0$; $var(\varepsilon_{i,L}) = var(\varepsilon_{i,H}) = 1$; $corr(\varepsilon_{i,L}, \varepsilon_{i,H}) = \rho$. Also, identification is achieved by assuming normality of the error terms (i.e. $\varepsilon_{i,L}$ and $\varepsilon_{i,H}$ are assumed to be distributed bivariate normal). A key condition for identification in this instrumental variables framework is that the instruments in equation (4.1) are uncorrelated with both $\varepsilon_{i,H}$ and $\varepsilon_{i,L}$ (Angrist & Pischke, 2009).

Given that there are only discrete measures of both LFP and health (i.e. SAH), the above model was implemented via a non-linear transformation of the linear index model in equations (4.1) - (4.2) as follows:

$$h_i = 1\left[X^/\beta^* + \gamma^* z_i + \varepsilon_{i,H} > 0\right] \qquad [4.3]$$

$$l_i = 1[X^/\alpha^* + \delta^* h_i^* + \varepsilon_{i,L} > 0] \qquad [4.4]$$

From equation (4.4), an individual is deemed to participate in the labour force if her underlying latent LFP index exceeds zero. An analogous interpretation holds for health status as shown in equation (4.3). Model (4.1)-(4.4) is similar to those used in modelling the impact of fertility on labour supply in the US (Angrist & Pischke, 2009) and the relationship between health status and poverty in South Africa (Godlonton & Keswell, 2005), where similar identification issues arise.

For the bivariate probit specification, the above model was estimated via full information maximum likelihood, a technique that yields efficient estimates. A formal test of exogeneity entails testing whether the correlation coefficient, ρ is statistically different from zero. If ρ is statistically significant, it is evidence of the endogeneity of SAH. Statistical insignificance of ρ on the other hand, would suggest that estimating separate LFP and SAH equations may not produce inconsistent estimates.

4.5.2 Data and empirical model specification

Most of the analysis in this chapter was based on the third (i.e. most recent) wave of the NIDS dataset. The sample was restricted to respondents aged between 20 and 60 years in wave 3 so as to exclude students and retirees from the analysis. Furthermore, Asians were excluded given their small sample size. Labour force participation was modelled on the extensive margin; therefore, the outcome is LFP, a dummy variable which equals one if the respondent is a labour force participant (broad definition, i.e. the respondent was willing to work in the past month) and zero otherwise. It is expected that a healthy individual will be more willing and able to work relative to one in poor health, ceteris paribus. Estimates from strict LFP models were used as robustness check on the estimates.

The main covariate of interest is SAH, a dummy variable derived from a five-level question regarding how an individual rates her health over the past month. Possible responses are: excellent, very good, good, fair and poor. In this analysis, healthy respondents (i.e. SAH=1) are those who self-reported their health to be excellent, very good or good, while those adjudged sick

(SAH=0) are those who declared fair or poor health. Though being "healthy" or "sick" is a complex phenomenon not fully captured by one's SAH response, these terms are adopted nonetheless for want of better terminology. As shown in Figure 4.2 below, LFP was not very different between the excellent, very good and good categories.

For the bivariate probit model, the empirical specification of model (4.1) - (4.4) is as follows:

$$\Pr(h_i = 1|X, Z) = \Phi\left(X^{/}\beta + \gamma z_i\right) \qquad [4.5]$$

$$\Pr(l_i = 1|X) = \Phi\left(X^{/}\alpha + \delta h_i\right) \qquad [4.6]$$

where h_i is SAH and l_i, observed LFP. For the general specification, X consists of household grant receipt, years of schooling, age dummies, location, race, provincial unemployment rate, marital status, gender, number of under-17 children in the household and household size. For the gender-specific models, X is identical to the above except for the exclusion of gender from both male and female equations and the inclusion of the presence of at least one employed male in the household in the female specification.

Z is a vector of SAH instruments. These are dummy variables indicating whether the respondent has experienced joint pain/arthritis and/or memory loss in the past 30 days. The approach used here hinges on the recognition that SAH is a summary/representative measure of overall health status, an assertion amply demonstrated in the literature (Benítez-Silva et al., 2004; Ferraro, 1980; LaRue, Bank, Jarvik, & Hetland, 1979; Nagi, 1969). This is also supported by the earlier point that SAH strongly predicts subsequent mortality. I assume that an individual who experiences say, joint pains would only change her labour market status due to illness only when such a condition so adversely affects her health that she rates it below "good". This informs the treatment of such health conditions as SAH instruments given that much as they directly affect health status (proxied by SAH), I see no compelling reason to directly include them in a LFP equation which has SAH as a covariate. Thus, I assume that they affect LFP only indirectly through their influence on overall health status (captured by SAH). This is similar to Bound's (1991) argument that merely proxying health status with relatively objective health measures in a labour supply equation does not solve the problem of endogeneity as objective health measures do not perfectly reflect work capacity. But I am not unmindful of the fact that regarding SAH responses as representing the totality of true health status may come across as a strong assumption. Unfortunately, this is as much as I can do in the present circumstance.

A potential argument against the suitability of these instruments is that unobserved individual characteristics associated with low LFP probability (e.g. low innate ability and drive) may also lead to a higher probability of having such health conditions (i.e. the instruments). To the best of my knowledge, there is no convincing evidence in the literature in this regard. Moreover, I regressed each of the instruments on parental education and parental mortality before the respondent reached five years (since these variables may affect a respondent's future labour market outcomes) as well as relevant covariates like own education, location and other socio-economic variables. Parental education and parental mortality were not statistically significant even at the 10% level for both males and females (results available on request).

Moreover, though I argued earlier that individuals are assumed to be rational when making self-assessments of their health status and that SAH may be considered to represent the totality of one's health, it may also be argued that the instruments may directly influence participation in specific sectors, e.g. joint pain/arthritis influencing manual jobs. Unfortunately, this could not be reliably tested due to data constraints but should be borne in mind when interpreting the results.

Furthermore, these instruments are considered largely exogenous to one's labour market status, i.e. they are not likely to be the result of rationalization of labour market status even though they are self-reported

(Au et al., 2005; Bound, 1991; Bound et al., 1995; Cai, 2010; Cai & Kalb, 2006; Stern, 1989). This is important as their validity will be questionable if they are determined by one's labour market status. For instance, if being employed under hazardous working conditions significantly determines any of the instruments, or if unemployed respondents who do not actually suffer from any of the conditions implied by the instruments declare that they suffer from them due to the shame/stigma associated with being unemployed or a non-labour force participant, the instruments may no longer be valid. To empirically test this, I regressed each instrument on the employment status dummy and a host of covariates like education, age, race and gender. The employment dummy was not statistically significant even at the 10% level in each of the regressions (results also available on request). This empirically strengthens the case for the exogeneity of the instruments. Finally, it is hoped that the array of socio-economic controls included in the health and LFP equations above helped purge the error term of most of the plausible reasons why it may be conditionally correlated with the instruments.

Earlier, I noted Bound's (1991) concern that merely instrumenting SAH with more objective health conditions may not be enough to identify the impact of health on labour supply if the measurement error associated with SAH is systematically correlated with, say, financial gains due to non-participation. In this case, the coefficient of SAH in the IV framework above over-estimates the effect of health on LFP. This is however not likely a significant issue in this study, as disability grant in South Africa is not predicated on non-labour force participation. But as earlier noted, the estimates are consistent if rationalization endogeneity is due to, say, tastes for leisure (see e.g. Au et al., 2005). I return to this point in the results section.

Descriptive statistics

Table 4.1 shows descriptive statistics of variables employed in the analysis.

Table 4.1: Descriptive statistics

Variable	N	Mean	Std.Dev.[a]
labour force participation	11250	0.73	0.4
self-assessed health	9822	0.90	0.3
matric	11271	0.37	0.5
age20 - 25	11294	0.22	0.4
age26 - 30	11294	0.15	0.4
age31 - 35	11294	0.14	0.3
age36 - 40	11294	0.13	0.3
age41 - 45	11294	0.11	0.3
age46 - 50	11294	0.10	0.3
age51 - 60	11294	0.15	0.4
traditional authority	11264	0.29	0.5
rural formal	11264	0.06	0.2
urban formal	11264	0.53	0.5
urban informal	11264	0.12	0.3
African	11294	0.83	0.4
coloured	11294	0.09	0.3
white	11294	0.08	0.3
married/cohabiting	9824	0.40	0.5
number of under-17 children in household	11294	1.43	1.7
male	11264	0.48	0.5
grant	11264	0.54	0.5
household has employed male	11285	0.50	0.5
provincial unemployment rate	11294	25.47	3.5
household size	11264	4.80	3.2
joint pain	9823	0.08	0.3
memory loss	9819	0.05	0.2

Source: NIDS wave 3; author's calculations; sample corrected for complex survey design and non-random attrition; [a] standard deviation

With regard to a priori expectations, education is expected to increase the probability of LFP. Also, it is likely to improve health through greater awareness of health-related knowledge (Cai & Kalb, 2006). Age is also likely to be positively correlated with LFP as older respondents tend to have more networks which is an important determinant of continued participation in the labour market. However, it has been found to be negatively associated with health status (see e.g. Cai & Kalb, 2006; Kenkel, 1995). Apart from the point made in Chapter 2 of a potential relationship between ones' location and race on the one hand and one's LFP status on the other, location and race also capture place and racial heterogeneity in health status. Compared to Africans and coloureds, whites are expected to have higher LFP owing to historical realities in the country. Also given that they are more likely to reside in urban formal centres with better medical facilities, whites are also expected to enjoy better health status. Evidence of a significant relationship between marital status and LFP exists in the literature, where the relationship is especially negative for females (Jaumotte, 2003). On the other hand, health is mainly seen to be positively associated with marital status (Beckett & Elliott, 2002). Grant receipt is likely to be negatively associated with LFP as non-labour income may result in higher reservation wage or early retirement (Mastrobuoni, 2009); conversely, non-labour income of the elderly has been associated with increased job search of prime age adults due to the extra income it brings to the household (Ardington, Case, & Hosegood, 2009). Provincial unemployment rate is expected to be negatively associated with LFP (Dinkelman & Pirouz, 2002; Evans & McCormick, 1994). This variable is used as a proxy for local labour market conditions, an attempt to capture some demand-side determinants of labour supply. Having an employed male (usually a primary breadwinner especially in an African context) may put less economic pressure on the woman, thereby resulting in reduced female LFP especially if the hypothesis that women are the main producers of domestic services like childbearing and child care holds (Joll, McKenna, McNabb, & Shorey, 1983); but it may also lead to a rise in female LFP due to increased labour market-related information/networks (Dinkelman & Pirouz, 2002).

Table 4.1 shows that 73% of the sample were labour force participants while 90% reported being in the excellent, very good or good health category. Married/cohabiting respondents made up 40% of the sample, while 48% were male. With regard to the SAH instruments, 8% and 5% suffered from joint pain/arthritis and memory loss respectively.

There appears to be a negative relationship between age and health status as older respondents were more likely to be in worse SAH categories and less likely to be in healthy categories relative to their younger counterparts as shown in Table 4.2 below. Furthermore as shown in Table 4.3, fair and poor health were concentrated more among women, while men had higher proportions in the best two health categories. Therefore, there is tentative evidence of a negative relationship between age and being female on the one hand, and health on the other.

Table 4.2: SAH across the age distribution (row percentages sum to 100)

Age categories	Excellent	V.good	Good (%)	Fair	Poor	**Row total (N)**
20-25	43.9	31.1	21.9	2.5	0.6	**2374**
26-30	39.9	33.4	22.2	4.0	0.6	**1481**
31-35	40.1	27.5	23.7	5.6	3.1	**1138**
36-40	35.5	30.5	25.8	5.5	2.7	**1070**
41-45	29.0	30.1	28.4	9.0	3.5	**978**
46-50	23.9	31.9	28.6	12.0	3.5	**1063**
51-60	19.4	25.3	32.5	16.1	6.7	**1718**
Total	**34.0**	**29.9**	**25.8**	**7.4**	**2.8**	**9822**

Source: NIDS wave 3; author's calculations; sample corrected for complex survey design and non-random attrition

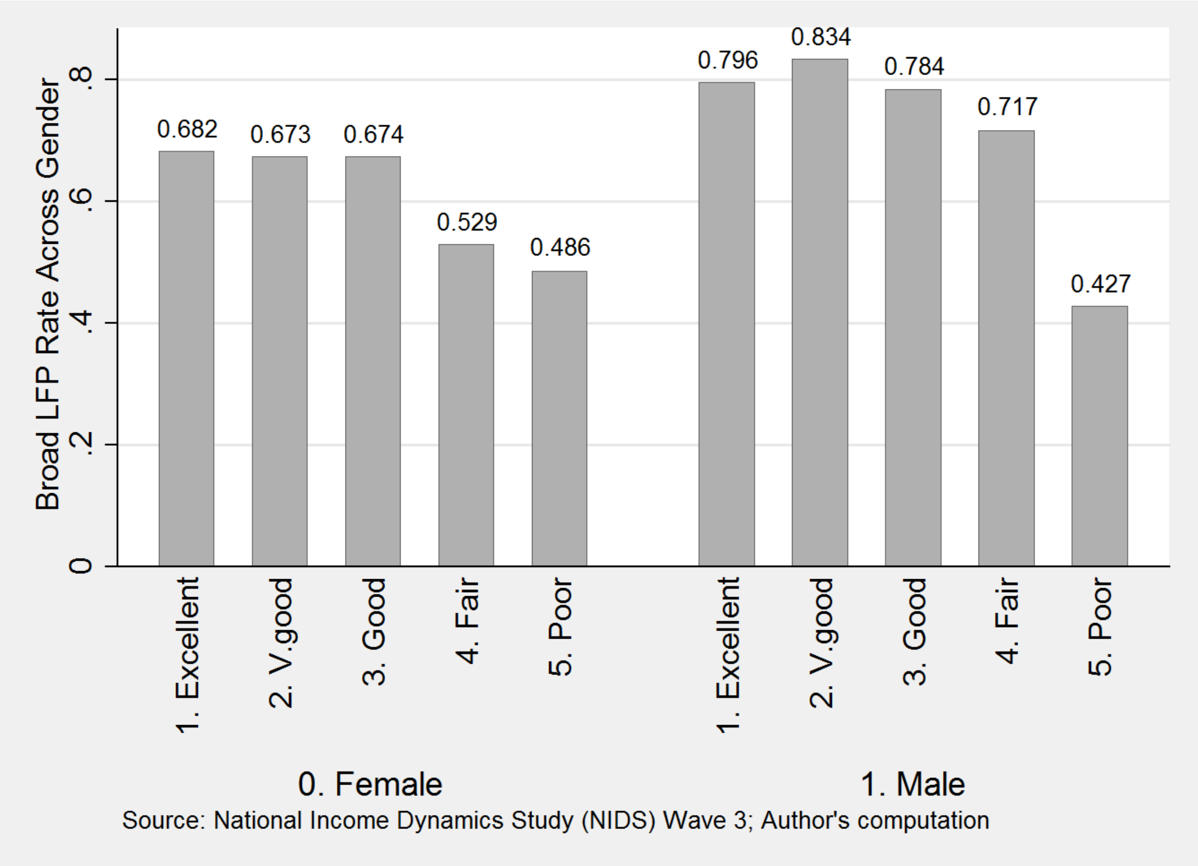
Table 4.3: SAH across gender groups (row percentages sum to 100)

	SAH categories (%)					
	Excellent	Very good	Good	Fair	Poor	**Row total (N)**
Female	30.9	29.5	27.2	8.9	3.6	**5858**
Male	37. 8	30.5	24.2	5.8	1.8	**3964**
Total	**34.0**	**29.9**	**25.8**	**7.4**	**2.8**	**9822**

Source: NIDS wave 3; author's calculations; sample corrected for complex survey design and non-random attrition

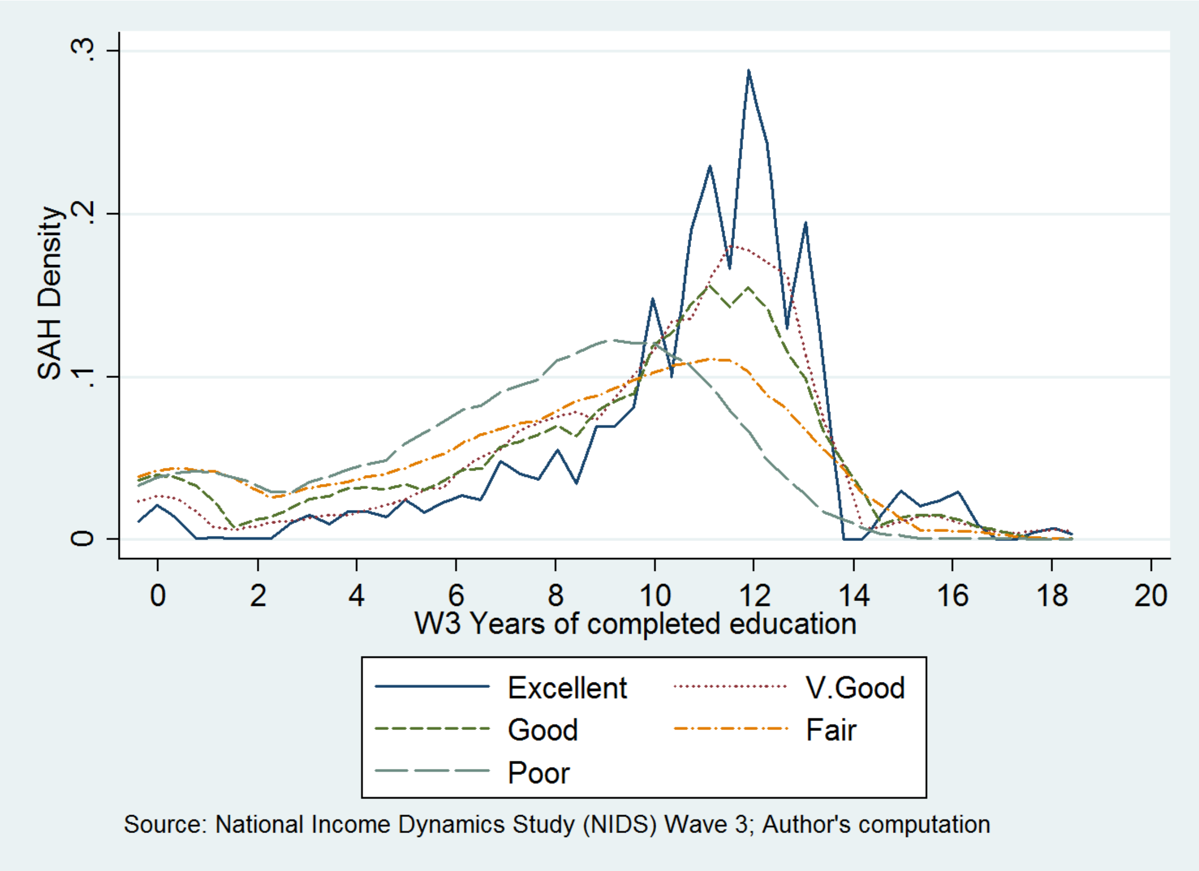
Figure 4.2 below reveals a positive relationship between better health status and LFP for both males and females. Apparently, there was not much heterogeneity in the LFP rates of the various groups classified as healthy. Furthermore, males had higher participation rates across all SAH categories (except poor) than females. Between-gender (i.e. male-female) percentage point differences ranged from -6 for the poor health category to 19 for the fair category. Figure 4.2 also perhaps highlights structural features of the South African labour market unrelated to health, as broad LFP rate among men in fair health exceeded that of women in even excellent health.

Figure 4.2: Health-LFP relationship by gender



With regard to health and human capital accumulation, Figure 4.3 shows relatively higher densities of poor health outcomes compared to better outcomes at lower educational levels while the converse obtained at higher educational levels. This lends credence to the hypothesis that healthier individuals are more likely to acquire education.

Figure 4.3: Distribution of SAH by years of schooling



In line with the use of a two-category health variable in the regressions, Table 4.4 reveals that the proportion of non-labour force participants who self-reported being sick (i.e. fair or poor health) was twice that of labour force participants for both strict and broad LFP. This is a tentative indication of a positive relationship between better SAH and LFP.

Table 4.4: Health-LFP status in wave 3 (row percentages sum to 100)

	Health Status (%)		**Row total (N)**
LFP Status	Sick	Healthy	
Strict Non-LFP	15.8	84.2	**3848**
Strict LFP	7.8	92.2	**6318**
Broad Non-LFP	16.0	84.0	**3217**
Broad LFP	8.0	92.0	**6585**

Source: NIDS; author's calculations; sample corrected for complex survey design and non-random attrition

Finally, I consider a descriptive relationship between SAH and its instruments in Table 4.5.

Table 4.5: Distribution of instruments across SAH categories

		SAH categories (%)		Row size
	Has condition?	Sick	Healthy	
joint pain/arthritis	No	7.7	92.3	**8949**
	Yes	40.2	59.8	**870**
Memory loss	No	9.0	91.0	**9295**
	Yes	35.0	65.0	**520**

Source: NIDS wave 3; author's calculations; sample corrected for complex survey design and non-random attrition

Table 4.5 provides preliminary evidence of non-trivial correlation between SAH and the instruments. For instance among those with no joint pain/arthritis, only 8% reported being in fair or poor health while it was 40% of those with the condition. A similar finding was obtained for memory loss.

The foregoing descriptive analysis shows substantial scope for the incorporation of health into LFP determination. It has been shown that non-labour force participants were twice as likely as participants to report worse SAH (Table 4.4). Also, other relationships such as that between education and LFP might have a health undertone as health is an important determinant of the amount of education acquired by an individual. Ascertaining the health-LFP relationship is the focus of the following analysis.

4.5.3 Results and discussion

As a preliminary analysis, Table A4.1 in the appendix depicts a traditional LFP model and the results generally conformed to a priori expectations. For instance, education, age, being male and location (relative to living in traditional authority areas) were associated with increased LFP probability. Marriage/cohabitation, which was associated with higher male participation had no relationship with female LFP. Receipt of government grant was associated with reduced LFP as expected while the presence of at least one employed male in the household was associated with increased female participation, apparently suggesting positive network externalities. A surprising result is the statistical insignificance of race. This did not change even with the exclusion of location which is a possible confounder given the strong link between race and location discussed in Chapter 2. As shown shortly, these results are very similar to what was obtained after controlling for health. Therefore, the discussion only focuses on the latter.

For the main analysis, attention is focused on the bivariate probit and IV-LPM models in Table 4.6. In line with Angrist and Pischke (2009), a useful way to think about policy-relevant measures in the bivariate probit framework is to consider marginal effects rather than difficult-to-interpret index coefficients obtainable from model 4.5 - 4.6 above.

Table 4.6: Marginal effects of SAH and controls on LFP probability (bivariate probit and IV-LPM)

	(1)	(2)	(3)	(4)	(5)	(6)
		Bivariate probit			IV-LPM	
Variables	Total	Female	Male	Total	Female	Male
sah	0.22***	0.20*	0.25*	0.22**	0.17	0.19
	(0.07)	(0.11)	(0.14)	(0.08)	(0.11)	(0.13)
grant	-0.06***	-0.05***	-0.07***	-0.07***	-0.06***	-0.07***
	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.03)
matric	0.10***	0.15***	0.05*	0.10***	0.15***	0.05*
	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.03)
age26-30	0.14***	0.18***	0.09***	0.15***	0.19***	0.11***
	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)	(0.03)
age31-35	0.15***	0.18***	0.12***	0.16***	0.19***	0.12***
	(0.02)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)
age36-40	0.19***	0.21***	0.15***	0.19***	0.22***	0.14***
	(0.02)	(0.03)	(0.04)	(0.02)	(0.03)	(0.03)
age41-45	0.15***	0.15***	0.15***	0.16***	0.16***	0.14***
	(0.02)	(0.03)	(0.05)	(0.03)	(0.03)	(0.04)
age46-50	0.11***	0.13***	0.05	0.11***	0.13***	0.06
	(0.02)	(0.03)	(0.04)	(0.03)	(0.04)	(0.04)
age51-60	0.01	0.02	-0.03	-0.00	-0.00	-0.03
	(0.03)	(0.04)	(0.04)	(0.03)	(0.04)	(0.04)
rural formal	0.07***	0.03	0.11***	0.09***	0.04	0.12***
	(0.03)	(0.03)	(0.04)	(0.03)	(0.04)	(0.03)
urban formal	0.11***	0.12***	0.08***	0.12***	0.13***	0.09***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)
urban informal	0.06**	0.09***	0.02	0.08**	0.11***	0.03
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)
African	0.03	0.05	-0.01	0.03	0.05	0.01
	(0.04)	(0.05)	(0.05)	(0.04)	(0.05)	(0.04)
coloured	0.00	0.06	-0.08	0.01	0.06	-0.06
	(0.05)	(0.06)	(0.06)	(0.04)	(0.05)	(0.06)
prov. unemp†	0.00	-0.00	0.00*	0.00	-0.00	0.00*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
married††	0.02	-0.04*	0.09***	0.02	-0.04*	0.08***
	(0.01)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)
male	0.10***			0.10***		
	(0.01)			(0.01)		
num. children‡	-0.00	-0.00	-0.00	-0.01	-0.01	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
household size	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)
employed male‡‡		0.04**			0.05**	
		(0.02)			(0.02)	
constant				0.26**	0.31**	0.40**
				(0.12)	(0.15)	(0.17)
R^2				0.12	0.12	0.11
F-stat	42.14	26.65	13.36	47.8	27.3	17.3
rho	-0.23	-0.20	-0.28			
	(0.14)	(0.19)	(0.30)			
N	9775	5825	3950	9775	5825	3950

Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1; estimates corrected for complex survey design and non-random attrition; †provincial unemployment rate; ††married/cohabiting; ‡number of under-17 children in household; ‡‡household has at least one employed male

Given these results from the bivariate probit model, Angrist and Pischke maintained that one can recover estimates of average treatment effect (ATE), average effect of treatment on the treated (TOT) and local average treatment effect (LATE) of the impact of health (viewed as a kind of endogenous treatment in this context) on LFP. On the other hand, the IV-LPM only yields the LATE estimator. Similar measures have been recovered from both the bivariate probit and IV-LPM models in a study of the effect of fertility on female employment where similar identification issues arise as in this study (Angrist & Pischke, 2009). Therefore, following Angrist and Pischke, ATE, i.e.

$$E\left[l_{1,i} - l_{0,i}\right] = E\left\{1\left[X^{/}\alpha^{*} + \delta^{*} > \varepsilon_{i,L}\right] - 1\left[X^{/}\alpha^{*} > \varepsilon_{i,L}\right]\right\}$$

can be specified as:

$$E\left\{1\left[X^{/}\alpha^{*} + \delta^{*} > \varepsilon_{i,L}\right] - 1\left[X^{/}\alpha^{*} > \varepsilon_{i,L}\right]\right\} = E\left\{\Phi\left[\frac{X^{/}\alpha^{*} + \delta^{*}}{\sigma_{\varepsilon_{i,L}}}\right] - \Phi\left[\frac{X^{/}\alpha^{*}}{\sigma_{\varepsilon_{i,i}}}\right]\right\} \quad [4.7]$$

while TOT, i.e. $E\left[l_{1,i} - l_{0,i} | h_i = 1\right]$ takes the following form:

$$E\left\{1\left[X^{/}\alpha^{*} + \delta^{*} > \varepsilon_{i,L}\right] - 1\left[X^{/}\alpha^{*} > \varepsilon_{i,L}\right] | X^{/}\beta^{*} + \gamma^{*} z_i > \varepsilon_{i,H}\right\}$$

$$= E\left\{\frac{\Phi_b\left(\frac{X^{/}\alpha^{*} + \delta^{*}}{\sigma_{\varepsilon_{i,L}}}, \frac{X^{/}\beta^{*} + \gamma^{*} z_i}{\sigma_{\varepsilon_{i,H}}}; \rho_{\varepsilon_{i,L}\varepsilon_{i,H}}\right) - \Phi_b\left(\frac{X^{/}\alpha^{*}}{\sigma_{\varepsilon_{i,L}}}, \frac{X^{/}\beta^{*} + \gamma^{*} z_i}{\sigma_{\varepsilon_{i,H}}}; \rho_{\varepsilon_{i,L}\varepsilon_{i,H}}\right)}{\Phi\left(\frac{X^{/}\beta^{*} + \gamma^{*} z_i}{\sigma_{\varepsilon_{i,H}}}\right)}\right\} \qquad [4.8]$$

Following Chiburis et al. (2011), the LATE estimator from the bivariate probit model is specified thus:

$$LATE$$

$$= \frac{\left[\Phi_b\left(X^{/}\beta^{*} + \gamma^{*}, X^{/}\alpha^{*} + \delta^{*}; \rho_{\varepsilon_{i,L}\varepsilon_{i,H}}\right) + \Phi_b\left(-\left(X^{/}\beta^{*} + \gamma^{*}\right), X^{/}\alpha^{*}; -\rho_{\varepsilon_{i,L}\varepsilon_{i,H}}\right)\right] - \left[\Phi_b\left(X^{/}\beta^{*}, X^{/}\alpha^{*} + \delta^{*}; \rho_{\varepsilon_{i,L}\varepsilon_{i,H}}\right) + \Phi_b\left(-X^{/}\beta^{*}, X^{/}\alpha^{*}; -\rho_{\varepsilon_{i,L}\varepsilon_{i,H}}\right)\right]}{\Phi\left(X^{/}\beta^{*} + \gamma^{*}\right) - \Phi\left(X^{/}\beta^{*}\right)} \qquad [4.9]$$

where l_0, l_1, Φ, Φ_b, σ and ρ are the LFP outcome of the non-treated, the LFP outcome of the treated, normal cumulative distribution function (CDF), bivariate normal CDF, error variance and error correlation coefficient respectively. The simultaneous estimation of the first and second stage equations is implicit in the above system.

Especially important for the LATE estimator in the above bivariate probit framework, I assume a homogenous effects model. This assumption also informs subsequent statistical tests with regard to the IV model. Though an apparently strong assumption, it is not as restrictive as it seems. This is because, though the LATE parameter capturing the impact of health on LFP (i.e. δ) appears as a constant effect in the above specification, it can be conceptualized as a heterogeneous effects coefficient obtained by estimating the LATE estimator for each X_i by IV and averaging using the histogram of covariates. This capability also makes the LATE estimator robust to violations of the normality assumption (Angrist & Pischke, 2009). For IV-LPM, the LATE estimator is obtained as the usual marginal effect in a two stage least squares framework.

Table 4.7 depicts ATE, LATE and TOT estimates of the impact of health on LFP obtained from the above bivariate probit model and IV-LPM. Standard errors for the bivariate probit model were obtained from 400 bootstrap replications.

Table 4.7: Effect of health on LFP

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Variables	Bivariate probit									IV-LPM marginal effects		
	ATE			TOT			LATE			LATE		
	Total	Female	Male	Total	Female	Male	Total	Female	Male	Total	Female	Male
sah	0.23***	0.20**	0.29*	0.26***	0.23**	0.33*	0.24***	0.20**	0.29*	0.22**	0.17	0.19
	(008)	(0.09)	(0.17)	(0.09)	(0.10)	(0.09)	(0.08)	(0.08)	(0.17)	(0.08)	(0.11)	(0.13)
N	9775	5825	3950	9775	5825	3950	9775	5825	3950	9775	5825	3950

Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1; estimates corrected for complex survey design and non-random attrition

From Table 4.7, ATE was 0.23 for the general (i.e. total) specification using equation (4.7). For the female and male specifications, ATE was 0.20 and 0.29 respectively. Using equation (4.8), TOT estimates were 0.26, 0.23 and 0.33 for the general, female and male specifications respectively. As indicated above, the set of bivariate probit estimates that should be directly comparable with IV-LPM estimates are the LATE estimates (columns 7-9). Apart from the male coefficient, both models were quite numerically similar, with bivariate probit coefficients slightly higher than their IV-LPM counterparts (Abadie (2000b) and Angrist (2001) also found higher bivariate probit estimates relative to linear IV estimates). Even the male bivariate probit LATE estimate that seemed to differ much from its IV-LPM counterpart was only statistically significant at 10%. The generally higher LATE estimates may be an indication that the effect of SAH on LFP among compliers (i.e. the sub-population whose memory loss and joint pain status affected their SAH ratings) was higher than in the general population. Table 4.7 suggests that depending on the estimator, being healthy resulted in increases in LFP probability of 20-23%, 29-33%, and 22-26% among the female, male and general populations respectively. These are similar to estimates of the impact of health (SAH) on employment (25% and 19%) obtained among older men and women (50-64 years) respectively in Canada (Au et al., 2005).

First stage estimates and instrument relevance

Full first stage results for the IV-LPM are presented in Table A4.2 in the appendix (covariates conformed to theoretical expectations and were mostly statistically significant at conventional levels) while the coefficients of only the instruments are presented in Table 4.8 so as to keep the discussion focused. The results clearly show that each of the instruments significantly predicted SAH even at the 1% level of significance. They also conformed to theoretical expectations, as suffering from any of the conditions was associated with increased probability of being in poor health relative to not suffering from it. The F statistics in a joint test of the instruments in all specifications (see Table 4.8) exceeded the Staiger and Stock (1997) critical F statistic of 10, implying that the instruments were not weak. Similar conclusion was reached for the gender-based models.

Table 4.8: First stage LPM estimates: marginal effects

	Dependent variable: $\Pr(sah = 1\|X)$		
	(1)	(2)	(3)
Variables	Total	Female	Male
joint pain	-0.25***	-0.25***	-0.23***
	(0.03)	(0.03)	(0.06)
memory loss	-0.17***	-0.15***	-0.20***
	(0.04)	(0.04)	(0.06)
N	9795	5836	3956
F (for joint pain & memory loss)	60.5	56.9	16.5
Hansen J (p value)	0.11	0.51	0.50

Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1; estimates corrected for complex survey design and non-random attrition;

Instrument validity

Given the assumption of a homogenous effects model, to formally ascertain instrument validity for the IV-LPM, Hansen's J test of over-identifying restrictions failed to reject the null hypothesis of valid instruments even at 10% as shown in Table 4.8.

With regard to whether relatively objective health conditions directly explained LFP, I included different relatively objective health measures (relative to SAH) in a LFP equation that excluded SAH[9]. The relatively objective health conditions include: serious injury, body ache, persistent cough, rash, painful urination, back ache and tight chest among others (results available on request). None of these other objective health controls was statistically significant even at the 10% level. This is expected and is in line with Bound et al. (1999) who maintained that these measures also suffer from measurement error in that they do not perfectly reflect work capacity (thus, making them susceptible to attenuation bias inherent in

[9] The bias caused by including the potentially endogenous SAH in this regression has been noted by Murray (2006).

poorly-measured variables when used as health proxies in a labour supply regression).

A potential concern regards the generalizability of the above bivariate probit and IV-LPM results over the entire population as it may be argued that the nature of the instruments makes them likely to disproportionately affect old individuals. To ascertain whether the above results were driven by old respondents, I re-estimated the models on a sub-sample younger than 50 years. Across all estimators (ATE, TOT and LATE), the results were very similar to the above estimates, mainly differing in the second decimal place. Also, the patterns of statistical significance were similar. The same is true of the IV-LPM results. These results are available on request.

With regard to the controls in Table 4.6 above, the marginal effect of interest for the kth continuous covariate in the LFP equation is $\frac{\partial \Pr(lfp=1|X_{-k})}{\partial X_k}$ (where X_{-k} are covariates other than X_k). This is the marginal effect of each covariate on the probability of a respondent being a labour force participant conditional on other covariates. For each discrete covariate, the change in LFP probability is in response to a unit change in such a covariate.

Table 4.6 (i.e. the health-augmented model) generally replicated the relationship between regression controls and LFP earlier found in the "traditional" specification (Table A4.1). A key qualitative difference was the significant negative association between marriage/cohabitation and LFP among women in the health-augmented model. Also, the highest age-LFP gradient now occurred later for men (mid-forties) while the female relationship appeared smoother with the highest gradient in the late thirties. This relationship mirrors findings in Ntuli and Wittenberg (2013) where only the oldest age cohort (in their case, aged 55-59 years) did not have significantly higher LFP relative to 20-24 year old African women. Average marginal effects of regression controls from both specifications (i.e. bivariate probit and IV-LPM) were almost identical and the estimates largely conformed to a priori expectations and were statistically significant at conventional levels. For instance, household grant receipt was associated with 5-7% reduction in LFP probability across both models, a finding similar to Bertrand et al.'s (2003) result of 4% reduction of employment probability associated with household eligibility for the old age pension among Africans. Also, having at least a matric was associated with 5-15% increase in LFP probability. Furthermore, spatial distribution mattered for LFP (apparently reflecting persistent effects of apartheid-era living arrangements where most jobs were located away from informal areas mostly occupied by non-whites); LFP probability was generally higher in other locations relative to traditional authority areas mainly populated by Africans. Local labour market conditions (proxied by provincial unemployment rate, which is a proxy for demand-side determinants of LFP) generally had no association with LFP. Marriage/cohabitation was negatively (positively) associated with female (male) participation, thereby supporting the theory of marriage and labour market participation which suggests that single women are more likely to be economically active compared to their married counterparts while the converse holds for men (Becker, 1981). Negative marital status-LFP relationship for women has also been documented in South Africa (Ntuli & Wittenberg, 2013; Posel & Casale, 2003). Males had 10% higher LFP rates than females. For females, the positive sign on the presence of at least one employed male in the household negates the added worker effect (Jaumotte, 2003), suggesting that male participation was associated with increased female participation. This likely captures network effects as suggested by Dinkelman and Pirouz (2002). The flip side is that families without male participants may become poorer, as not having at least one employed male in the household was associated with 4-5% decline in the probability of female participation.

Results in Table 4.6 (for both SAH and controls) were robust to the outcome being strict LFP except that the female SAH coefficient was no longer statistically significant at conventional levels. These results are available on request.

Sensitivity analysis

The following sensitivity analyses were conducted to ascertain the robustness of the above estimates.

Disability grant exclusion

It was earlier noted that systematic over-reporting of illness by non-labour force participants due to the expectation of financial reward (e.g. disability-based grants for non-participation) may render both estimates of SAH and such financial variables inconsistent even when objective health measures are used to instrument for SAH (see e.g. Au et al., 2005). Therefore, I stripped the household grant receipt variable of its disability grant component. Though the disability grant scheme in South Africa is not predicated on non-labour force participation per se, it is the component of the government's grant basket that *directly* targets economic agents with some kind of incapacity and may be the most likely economic reason (however remote) why respondents may give spurious health assessments. The results (available on request) showed virtually identical results for both SAH and the other controls, while the grant variable had slightly lower coefficients across specifications (but remained negatively statistically significant). Thus, it is not likely that the results in Table 4.6 and Table 4.7 were over-estimates of the impact of health due to rationalizing non-participation in the labour market in expectation of illness-related financial reward in the form of disability grants.

Testing for exogeneity of SAH

It was earlier noted that the dominant view in the literature is that self-reported health is an endogenous determinant of LFP, necessitating the use of instrumental variables as done in the foregoing analysis. However, it was also noted that some previous studies could not reject the hypothesis of exogenous SAH. If SAH is truly an exogenous determinant of LFP in any empirical context, applying IV techniques may be superfluous and will increase the likelihood of failing to reject the null hypothesis of no significant health impact when such impact might truly exist, given the relative inefficiency of IV estimates compared to ordinary least squares (OLS) (Cameron & Trivedi, 2010; Murray, 2006).

A formal test of SAH exogeneity in the bivariate probit framework is a Wald test of the statistical significance of the cross-equation correlation (ρ) between the LFP and SAH equations. If ρ is statistically different from zero, there is evidence of SAH endogeneity and vice versa. In this case, a positive ρ suggests that unobserved determinants of LFP are positively associated with unobserved determinants of SAH while the converse holds for a negative ρ (see e.g. Cai & Kalb, 2006; Stern, 1989). Table 4.6 above suggests failure to reject the null hypothesis of SAH exogeneity, i.e. $\rho = 0$, across all specifications. Also, Durbin-Wu-Hausman tests of the null hypothesis that SAH is exogenous failed to reject the null hypothesis of SAH exogeneity in all three specifications given p values of 0.2, 0.5 and 0.7 for the total, female and male specifications respectively. These suggest that SAH may not be endogenous to LFP in this context. The above conclusion of no SAH endogeneity is not novel as Stern (1989) found evidence of weakly endogenous self-reported disability on LFP in the US. Indeed, Bound et al. (1999) observed that most of the literature on the effect of health on labour force behaviour treats health as exogenous. On the contrary, some studies have rejected the hypothesis of exogeneity of SAH in Australia (Cai, 2010; Cai & Kalb, 2006). Thus, while SAH might prove to be exogenous in certain contexts, such conclusions appear to be context-specific and should be informed by appropriate statistical tests.

Marginal effects from a regression of LFP on SAH under the assumption of SAH exogeneity are shown in Table 4.9 below. Comparing these estimates for the regression controls with those in Table 4.6, one observes that the models appear to be robust to the exogeneity assumption. Perhaps this reinforces the findings from the more formal Durbin-Wu-Hausman and error correlation tests above. However, there was a marked decline in the marginal effect of SAH especially in the probit model in Table 4.9 (relative to the bivariate probit marginal effects). Here, being “healthy” was associated with 10-11%, 12-14% and 10-12% increase in LFP probability in the female,

male and general populations respectively. However, the relative efficiency of single equation relative to IV estimates resulted in statistically significant SAH marginal effects across all specifications even with these numerically smaller estimates. In a similar study in Canada, Au et al. (2005) found smaller coefficient of SAH in an employment equation when SAH endogeneity was not accounted for, compared to IV estimates.

Table 4.9: LFP determination allowing for exogeneity of SAH (marginal effects)

	(1)	(2)	(3)	(4)	(5)	(6)
		Probit			LPM	
Variables	Total	Female	Male	Total	Female	Male
sah	0.10***	0.10***	0.12***	0.12***	0.11***	0.14***
	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)	(0.04)
grant	-0.07***	-0.06***	-0.08***	-0.07***	-0.06***	-0.08***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)
matric	0.11***	0.16***	0.05**	0.11***	0.16***	0.05**
	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.03)
age26_30	0.14***	0.18***	0.09***	0.15***	0.19***	0.11***
	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)	(0.03)
age31_35	0.15***	0.17***	0.11***	0.15***	0.18***	0.12***
	(0.02)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)
age36_40	0.19***	0.21***	0.15***	0.18***	0.22***	0.13***
	(0.02)	(0.03)	(0.04)	(0.02)	(0.03)	(0.03)
age41_45	0.14***	0.14***	0.13***	0.15***	0.15***	0.13***
	(0.02)	(0.03)	(0.05)	(0.02)	(0.03)	(0.04)
age46_50	0.09***	0.12***	0.04	0.10***	0.12***	0.06
	(0.02)	(0.03)	(0.04)	(0.03)	(0.03)	(0.04)
age51_60	-0.01	-0.00	-0.05*	-0.02	-0.02	-0.04
	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)
rural formal	0.07***	0.02	0.11***	0.08***	0.03	0.12***
	(0.03)	(0.03)	(0.04)	(0.03)	(0.04)	(0.03)
urban formal	0.11***	0.11***	0.08***	0.12***	0.12***	0.09***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)
urban informal	0.06**	0.09***	0.02	0.08**	0.10***	0.03
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)
African	0.03	0.04	-0.01	0.03	0.05	0.01
	(0.04)	(0.05)	(0.05)	(0.04)	(0.05)	(0.04)
coloured	-0.00	0.06	-0.08	0.01	0.06	-0.06
	(0.05)	(0.06)	(0.06)	(0.04)	(0.05)	(0.05)
prov. unemp†	0.00	-0.00	0.00*	0.00	-0.00	0.00*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
married††	0.02	-0.03	0.10***	0.02	-0.03*	0.08***
	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)
male3	0.11***			0.11***		
	(0.01)			(0.01)		
num. children‡	-0.00	-0.00	-0.00	-0.01	-0.01	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
household size	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)
employed male‡‡		0.04**			0.05**	
		(0.02)			(0.02)	
constant				0.36***	0.38***	0.44***
				(0.08)	(0.10)	(0.10)
R^2				0.13	0.12	0.11
F-Stat	36.3	20.1	12.9	50.0	28.0	18.1
N	9784	5830	3954	9784	5830	3954

Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1; estimates corrected for complex survey design and non-random attrition; †provincial unemployment rate; ††married/cohabiting; ‡number of under-17 children in household; ‡‡household has at least one employed male

Also, a regression of LFP on the five-category SAH showed that other categories (in particular) excellent, very good and good categories were associated with significantly higher LFP relative to poor health (results are available on request).

Past health versus current LFP

Another way to look at the relationship between health and LFP if one suspects endogeneity of SAH is to exploit the timing of SAH responses relative to LFP status. Given the panel nature of the dataset, one can estimate current LFP as a function of past SAH, an approach that has been adopted in the literature (see e.g. Bartel & Taubman, 1986; Haveman et al., 1994). Though this method does not completely solve the endogeneity problem, especially if past SAH responses were used to justify past LFP status, it does mitigate the contemporaneous feedback relationship that likely exists between both variables. Therefore, I regressed wave 3 LFP status on wave 1 SAH status, wave 1 LFP and the same controls used above. The marginal effects of past SAH are presented in Table 4.10 below (marginal effects of the controls are similar to what obtained in the main regressions and are therefore not reported; they are however available on request).The results (expectedly attenuated) show that health significantly affected LFP even four years after a given health assessment. The magnitude ranged from 6-9%.

Table 4.10: Relationship between past SAH and current LFP

	(1)	(2)	(3)	(4)	(5)	(6)
		Probit			LPM	
Variables	Total	Female	Male	Total	Female	Male
sah	0.06***	0.07**	0.09**	0.07***	0.07**	0.09**
	(0.02)	(0.03)	(0.04)	(0.02)	(0.03)	(0.04)
F statistic	38.4	29.8	17.4	56.4	29.8	17.4
N	8328	5217	3111	8328	5217	3111

Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1; estimates corrected for complex survey design and non-random attrition

Health shocks and labour force participation

Some studies have exploited health changes (otherwise termed health shocks) available in panel datasets to provide a plausible argument with regard to the labour supply effect of arguably exogenous variations in SAH (see e.g. Chirikos & Nestel, 1985; Gómez & Nicolás, 2006). Consequently, I estimated models for the effect of positive and negative self-reported health shocks on LFP. I categorized an individual as having suffered an adverse health shock if she identified herself as healthy in wave 1 but sick in wave 3 while the baseline category consisted of respondents who were healthy in both waves. On the other hand, an individual was identified as having experienced health improvement if she was sick in wave 1 but healthy in wave 3 while the benchmark category in this case comprised respondents who remained sick in both waves. For the first category, 7.6% of usable responses (total=7112 observations) recorded adverse health shocks among working age respondents in wave 3 while 71.9% of usable responses from the second category (total=1333 observations) recorded health improvement.

As Table 4.11 indicates, the results for adverse health shocks were qualitatively similar (in terms of sign) to what was obtained in the contemporaneous and between-wave analyses above: adverse health shocks were associated with a 12-14% decline in male LFP in wave 3 while such shocks were not associated with any change in female LFP. These estimates were however, smaller than the IV estimates in Table 4.6 and Table 4.7. The “total” coefficients representing 8-9% higher probability of inactivity due to adverse health shocks in the population are higher than Gomez and Nicholas’ (2006) estimate of 3-4% in Spain. Such differences may be due to the fact that Gomez and Nicholas estimated causal effects of adverse health shocks (which this study could not recover due to fewer data waves). They could also stem from apparent differences in levels of mechanization in both labour markets, as South Africa is a developing economy where labour market participation may be more strongly associated with health. Health improvement was statistically significant for both males and females. Labour force participation probability for those who experienced health improvement increased by 10-11%, 14-16% and 11% in the female, male and general populations respectively relative to those who remained in constant bad health. Thus, both health deterioration and health improvement were associated with higher percentage changes in LFP probability for men relative to women. This is consistent with all the evidence throughout this chapter.

Table 4.11: Health shocks and LFP (marginal effects)

	(1)	(2)	(3)	(4)	(5)	(6)
		Probit			LPM	
Variables	Total	Female	Male	Total	Female	Male
			Adverse health shock			
	-0.08***	-0.05	-0.12***	-0.09***	-0.06	-0.14***
	(0.03)	(0.03)	(0.04)	(0.03)	(0.04)	(0.05)
F statistic	26.5	14.1	11.4	34.6	18.1	14.1
N	7083	4318	2765	7083	4318	2765
			Health improvement			
	0.11***	0.10**	0.16***	0.11***	0.11**	0.14**
	(0.04)	(0.05)	(0.06)	(0.04)	(0.05)	(0.06)
F statistic	7.2	5.9	5.2	11.3	9.8	8.0
N	1327	947	380	1327	947	380

Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1; estimates corrected for complex survey design and non-random attrition

4.6 CONCLUSION

This chapter has ascertained the effect of health on labour force participation in South Africa. The foregoing discussion suggests that better health (measured as composite/global self-reported health) is positively related to LFP. When treated as endogenous in a contemporaneous setting, ATE, TOT and LATE estimates from the bivariate probit model suggest that SAH exerts positive effect on LFP on the general population as well as in the male and female sub-populations. These results show 20-23%, 29-33% and 23-26% increase in LFP probability among the “healthy” relative to the “sick” for the female, male and general populations respectively. As noted above, the higher LATE estimates may be indicative of LFP being more sensitive to health status among compliers than in other sub-populations. Moreover, given the rather strong assumption of a homogenous effect model, various diagnostic tests suggest that SAH may not be endogenous to LFP in South Africa, a result not inconsistent with previous cross sectional evidence in the US (Stern, 1989). Exploring dynamic health changes made possible by the panel nature of the data suggests that adverse health shocks are associated with significantly decreased LFP of around 12-14% and 8-9% in both the male and general populations respectively over the four-year period (2008-2012). Health improvement is associated with about 11% LFP increase among the female and general populations as well as 16% increase among males. This study has therefore shown that better health is positively related to labour force participation in South Africa and that the relationship is not only temporary.

CHAPTER 5

5 THE HEALTH-WAGE GRADIENT IN SOUTH AFRICA

5.1 INTRODUCTION

In Chapter 4, it was shown that better health has a positive effect on LFP (i.e. the willingness to participate in the labour market) in both cross-sectional and temporal settings in South Africa. Even if one finds the argument for causality tenuous, it was shown that a positive association exists. Beyond such an exercise, it is important to determine whether health is related to wages for the employed. As noted in Chapter 1, ascertaining the relationship between health and the labour market is important given that health constitutes human capital and that better health should enhance productivity while higher productivity and earnings engender welfare improvement (Becker, 1964; Jack & Lewis, 2009).

Though the argument that health should positively influence productivity and wages is intuitively appealing (Suhrcke et al., 2005), whether this relationship actually holds in a particular context as well as its magnitude remains an empirical question. Furthermore, such a relationship may depend on the particular health measure being analysed, where the health-wage/productivity gradient may differ by context and health condition. For instance, it has been found that different physical and mental health indicators like hypertension, heart disease, depression and other mental illnesses are associated with varying productivity penalties in the US (Goetzel et al., 2004). Additionally, such costs may vary across the earnings distribution even for a particular health indicator.

Though many studies have examined determinants of wages in South Africa, they mainly focused on factors like education, race, gender and occupational class with virtually none examining the potential relationship between health and wages (Bhorat, 2004; Lam, 1999; Mwabu & Schultz, 1996). Consequently, I examine the average and heterogeneous relationship between health and wages in South Africa. Furthermore, given the possibly nuanced relationship between health and wages, I examine the relationship between physical and psychological health on the one hand, and wages on the other. Though both physical and psychological health are very important in their own right and are likely to reinforce each other (Prince et al., 2007), it is important to ascertain the differential magnitudes of the wage penalty associated with poor physical and psychological health respectively, especially in the face of severely binding budget constraints typical of a developing country like South Africa. In the event that available resources are insufficient to carry out substantial interventions on both health measures, it may make economic sense to prioritize the measure which has a more pronounced relationship with productivity so as to possibly set off a virtuous cycle where increased output will eventually help achieve holistic health interventions. Finally, knowledge of the temporal dimension of the health-wage gradient is important as it will help ascertain whether the productivity cost of illness might be under-estimated in a cross sectional analysis.

5.2 RESEARCH AIM AND OBJECTIVES

The aim of this chapter is to ascertain the magnitude of the relationship between health and wages over the wage distribution in South Africa. Specifically, I intend to ascertain:

i) The magnitude of the gradient between physical health (proxied by the body mass index) and wages in South Africa for low, median and high earners.
ii) The magnitude of the gradient between psychological health/depression and wages in South Africa for low, median and high earners.
iii) Whether the gradients persist over time.

5.3 EMPIRICAL REVIEW OF THE HEALTH-WAGE/PRODUCTIVITY RELATIONSHIP

One … way to evaluate health indicators is to test them against other observable effects of health status…. Because we believe that healthier people are likely to be more productive and therefore, on average, better remunerated in their employment, it should be possible to validate or qualify the various yardsticks of health.
(Savedoff & Schultz, 2000, p. 7)

The above quote indicates that good health is expected to have a positive effect on productivity and earnings. Though many studies have been conducted in this regard and the results largely positive, there is no consensus on the magnitude of impact due to differences in methodology, health measures, etc. (Suhrcke et al., 2005).

Some prominent studies that have examined the health-productivity relationship employed randomized controlled trials (RCTs) to examine the impact of certain nutrients on productivity. For instance, Thomas and Frankenberg (2002) discussed a number of studies that examined the impact of iron on productivity. Generally, iron deficiency was negatively related to work capacity, endurance, energy efficiency and aerobic capacity (i.e. the maximum amount of oxygen that can be used by the body). In a study of the effects of iron on male rubber workers in Indonesia, it was shown that giving iron supplementation to workers with iron deficiency anaemia could raise such workers' output by approximately 20%, though selective attrition of those in the control group may have resulted in an over-estimation of the effect (Basta, Karyadi, & Scrimshaw, 1979). Also, an iron supplementation RCT conducted by Li et al. (1994) on Chinese cotton mill workers showed a 5% increase in energy efficiency among the treated group relative to the controls after 12 weeks of treatment. Treatment also resulted in a 17% increase in production efficiency though there was no increase in work output. Similarly, better health and economic success were observed among men who were offered a 120mg weekly iron supplement relative to those in the control group in the Work and Iron Status Evaluation study in

Indonesia (Thomas et al., 2003). Evidence also exists of the effects of other nutrients on wages. For instance, a high protein drink administered to children in their first three years of life in Guatemala increased male hourly wage rate by USD0.67 when they were aged 25-42 years (Behrman, 2009).

Most empirical studies of the health-wage relationship are based on observational studies resulting in correlations or causal relationships (the latter established using quasi-experimental methods). In the US for instance, Lee (1982) estimated a simultaneous equations model of the health-wage relationship in order to account for the endogeneity of both health and wages. The labour market outcome was log hourly wage, while the health variables were a dichotomous health limitation variable and SAH. Identification was achieved by excluding net family asset from the wage equation; and the square of experience, race and region dummies from the health equation. The validity of these identification restrictions have however been questioned (Currie & Madrian, 1999). Results showed 161-222% health effects on wages. In another US-based study, Chirikos and Nestel (1985) found significant wage reductions for whites with improving, deteriorating and constant poor health conditions relative to those in constant good health. However, there was no significant wage change for the black sub-samples across different health variables. In another instrumental variables study that exploited variations in the birth weight of monozygotic twins, it was found that increasing the birth weight of children in the lower ends of the birth weight distribution to the average US birth weight could increase future earnings by up to 26% (Behrman & Rosenzweig, 2004).

Jackle and Himmler (2010) estimated the impact of health on wages in Germany using a panel framework with corrections for sample selection and endogeneity. Data was from the 1995-2006 annual German Socio-Economic Panel. Number of doctor visits in the past 3 months was used to instrument SAH given its possible endogeneity to work participation. Results showed that the effect of health on participation was greater for IV participation models than their single equation counterparts for both men and women, while there was also evidence of sample selection bias. Also, health was a significant determinant of male wages only, while it generally had no significant effect on women.

Savedoff and Schultz (2000) surveyed considerable literature on the impact of health on productivity in Latin America. A key finding that cut across all the studies is that instrumenting for health resulted in higher health coefficients in wage regressions relative to OLS, a finding consistent with attenuation bias of OLS in the presence of possibly mis-measured health variables (see e.g. Cameron & Trivedi, 2005). To achieve identification, they suggested aggregating individual responses to questions regarding access to health services, or price of health inputs at, say, the municipal level. Such instruments include average distance from a household to the nearest clinic, community-level use of health and sanitation services, proportion of the community with access to potable water, and the number of medical personnel in a community. In other instances, government records on climate, geography and infrastructure for each community can be matched to specific individuals who were resident in such places when some significant events (like drought) occurred so as to identify the impact of such exogenous events on their current health condition. However, the underlying assumption is that place of residence is not related to average health heterogeneity, migration decisions, individual preferences affecting health, location of health and other service infrastructure, among others, an assumption that may be difficult to justify (Rosenzweig & Wolpin, 1986; Schultz, 1988). Furthermore, selection into wage employment should be accounted for where possible as it enhances the generalizability of the results. Studies in this regard include Murrugarra and Valdivia (1999) and Cortez (1999)- Peru (health indicator: number of days ill); Parker (1999)- Mexico (health indicators: disability, disabled days, self-reported relative health status and functional limitations); and Ribero and Nunez (2001)- Colombia (health indicators: disability and number of disabled days).

Also using community-level price variables as health instruments, Knaul (2001) examined the impact of age at menarche on female wages in Mexico. The study found that a one-year decrease in women's age at menarche was associated with a 23-26% increase in hourly wages on the average. However, the explanatory power of the instruments on age at menarche was somewhat weak and the data had no information on place of birth and migration. As a result, it was not possible to adequately control for the disease/health environment when the women were growing up (which is a critical time for nutritional investments).

Another strand of the literature has examined the relationship between anthropometric measures like height and BMI on wages and productivity. For instance, Schultz (2002) examined wage gains associated with height as a form of health human capital in Ghana, Brazil and the US. Instruments for height were community supply of health-related services and infrastructure (a proxy for local price of health), parental education and occupation, a combination of these two sets of instruments and ethnic/racial heterogeneity. IV estimates showed that height significantly and positively explained variations in log wage in general and that IV estimates exceeded OLS estimates.

There are problems in using height as a health measure capturing the effect of human capital in a wage regression. A component of height is genetically determined and may not reflect nutrient intake, which is important for increased productivity. Additionally, height evolves over a long period of time and partly reflects conditions during early childhood including investments made by parents. Therefore, its explanatory power in a wage regression is expected to decrease with the inclusion of socio-economic status indicators.

But, unlike height, BMI captures both short and long term aspects of nutrient intake (Thomas & Frankenberg, 2002). Though obesity is viewed as a problem especially in developed countries, it has been suggested that employers view higher BMI as a marker of good health in developing countries especially given the difficulty in observing true health, hence its positive association with wages (Foster & Rosenzweig, 1992; Foster & Rosenzweig, 1994). For instance, BMI was found to positively affect the wages of urban Brazilian employees and the self-employed (Strauss, 1986) while a 1% increase in BMI was associated with 2.7% increase in wages in Ethiopia (Croppenstedt & Muller, 2000).

However, a number of studies conducted on the impact of obesity on wages in developed countries found negative or no effect. Using the National Longitudinal Survey of Youth in the US, Register and Williams (1990) found that obesity reduced female wages by 12% but had no impact on male wages. Using the same data, Loh (1993) found no significant effect of obesity on wage *levels* for both men and women. However, the study found that obesity significantly affected wage *changes* negatively for males but had no significant effect for females. Also, examining the effect of obesity on wages in the US, Baum and Ford (2004) showed the existence of persistent obesity-induced wage penalty in the first two decades of employment for both men and women.

Other studies have examined the impact of past health status/sickness history on current wages. Examining the effect of sickness on earnings in Sweden, Andren and Palmer (2001) found that individuals who were sick in previous years had lower earnings than others even if the previously sick were currently healthy. The effects were more pronounced for men relative to women. In a detailed study of the life cycle profile of the health-wage relationship, Pelkowski and Berger (2004) examined labour market effects of temporary and permanent health conditions. They found that permanent health conditions resulted in larger negative wage reductions for females relative to males (males experienced larger declines of hours worked relative to females). Also, adverse negative labour market consequences for males resulted from the onset of health problems in the 40s, while the labour market effects of negative health problems for females peaked in their 30s on the average.

Some studies have investigated the impact of psychological/mental health on earnings/income. To the best of my knowledge, evidence in this regard with respect to Africa in particular and developing countries in general is lacking. Ettner et al. (1997) examined the relationship between psychiatric disorders and earnings. The earnings equation was only estimated for those with positive income, implying that possible selection into work was not accounted for. Results of both OLS and IV models showed that the presence of any psychiatric disorder had detrimental effects on income for women, while IV estimates for men were not significant at conventional levels. Also, IV estimates were generally higher than OLS results, apparently due to the attenuation bias typical of the latter. Moreover, the authors suspected attenuated estimates as the wage proxy was annual personal income (which included disability-related income). However, using fixed and random effects IVs on the British Household Panel, Contoyannis and Rice (2001) found that reduced psychological health resulted in a decline in wages for males while excellent SAH raised female wages. Also, health variables were positively correlated with time-invariant individual effect. In another study, mental health has also been associated with greater earnings penalty relative to physical ailment for disabled UK men and women (Jones et al., 2006). Serious mental illness (computed using the Composite International Diagnostic Interview of the WHO) has been associated with annual earnings reduction of about USD16300 twelve months after mental health examination in the US (Kessler et al., 2008). To obtain more exogenous estimates, another study found that psychological conditions in childhood resulted in permanently low earnings of about USD4094 per year in the US (Smith & Smith, 2010). Also, respondents who suffered serious mental illness earned a third less than median earnings in a WHO survey involving 19 countries in the Americas, Europe, the Middle East, Africa, Asia and New Zealand (Levinson et al., 2010). Using quantile regression, it was found that while average mental health effects on wages were not sizeable, substantial effects were present at the lower part of the wage distribution especially among women in the US (Marcotte & Wilcox-Gok, 2003).

Though some studies have been conducted on wage determination in South Africa, they did not ascertain the role of health. Mwabu and Schultz (1996) have examined the returns to education at various quantiles of the wage distribution using the 1993 Project for Statistics on Living Standards and

Development data (SALDRU, 1994). Usual controls in a standard Mincerian (Mincer, 1974) wage regression such as education, quadratic labour market experience and its square, as well as place of residence were included in the quantile wage regression. The results showed that private wage returns to schooling were twice as high for non-whites than for whites in 1993, coupled with the fact that the returns to schooling increased with educational level for both racial groups. Also, Mwabu and Schultz (2000) have observed that returns to secondary and tertiary education are higher for women than men in South Africa.

Bhorat and Leibbrandt (1999b) found that returns to secondary education exceeded those of primary and tertiary education for Africans. However, Chamberlain and van der Berg (2002) cautioned that estimates of returns to education would likely bias the impact of education on earnings if one does not adjust for educational quality. For both whites and blacks, higher quality adjustment factors (i.e. assuming that most of the years of schooling were of poor quality) tended to depress the returns to primary and secondary education and vice versa, while the converse obtained for tertiary education.

In pointing out the so-called feminization of the South African labour market between 1995 and 1999, Casale and Posel (2002) noted that while more women were entering the labour force relative to men, women experienced higher unemployment and those who secured employment were mostly employed in low-paying jobs, suggesting a gender-based wage penalty against women.

From the foregoing, it is evident that studies investigating wage determination in South Africa have not accounted for the potential effect of health. Also, most international studies have only focused on average health-wage relationships without investigating wage determination over the wage distribution. Furthermore, though some international studies have investigated the effect of psychological/mental health on wages, the results are mixed. Therefore in the following sections, I investigate wage determination over the wage distribution, accounting for possible censoring of wages as well as ascertaining the relationship between health (physical, psychological and general) and wages.

In ascertaining the relationship between health and wages, the potential endogeneity of health is apparent. Factors responsible for such an endogenous relationship include the omission of important but unmeasured wage determinants like ability, as well as bidirectional causality (given that higher wages can facilitate the purchase of better health care). Additionally, health measurement error, which may result from say, the poor under-reporting their health due to their inability to "afford" to be ill (McIntyre, Gilson, Valentine, & Soderlund, 1998) may plague observed health measures. Also, environmental (i.e. non-genetic) factors are known to influence health (Cawley, 2004). These factors are likely to imbue health with endogeneity in a wage regression. However, given the lack of credible instruments in the dataset, this study does not investigate causal gradients. This is surely a limitation of this study.

5.4 ANALYTICAL METHODS AND DESCRIPTIVE ANALYSIS

5.4.1 Theoretical model specification

From the foregoing review, it is apparent that a large chunk of the empirical work on the health-wage relationship examined the average relationship. However, a common observation is that a summary measure (say, the average) may be inadequate to fully characterize the wage distribution; for instance, it has been noted that upper earnings quantiles have been increasing while lower quantiles have been decreasing in the US (Angrist & Pischke, 2009). Even more relevant for the purpose of this chapter is that the relationship between a variable of interest and wages may differ across the wage distribution. For instance, as shown in the following descriptive analysis, the relationships between self-reported health status and other key covariates like education on the one hand, and wages on the other, appear to differ at different parts of the earnings distribution. Quantile regression therefore represents a powerful tool with which to model such apparent non-linearities in the relationship of interest as it can help uncover the heterogeneous relationship between the outcome and covariate of interest, a relationship that may be masked when examining the average gradient. In this case, the coefficients of the various characteristics at each quantile represent rates of return to these characteristics at the specified quantile of the conditional wage distribution (Melly, 2005), quantities that cannot be obtained with OLS regression. Other advantages of this estimation strategy relative to OLS include its robustness to heteroscedasticity, outliers and non-normality of the error term, which are problems that often plague the typical wage distribution (Cameron & Trivedi, 2005; 2010; Powell, 1986a). Given that a nontrivial proportion of the sample was censored (i.e. their characteristics were observed while they did not have information on wages given that they were unemployed) – indeed 56% and 44% of waves 1 and 3 samples respectively were left-censored- I estimated censored quantile wage regressions as espoused by Powell (1986a) and Cameron and Trivedi (2005).

Generally for a linear continuous function,y, the qth quantile regression estimator, $\widehat{\beta_q}$ minimizes the loss function:

$$Q_N(\beta_q) = \sum_{i:y_i \geq x_i^{/}\beta}^{N} q|y_i - x_i^{/}\beta_q| + \sum_{i:y_i < x_i^{/}\beta}^{N} (1-q)|y_i - x_i^{/}\beta_q| \qquad [5.1]$$

over β_q, where x is a vector of regressors and β_q indicates that different quantiles estimate different values of β.

The starting point of Powell's (1986a) estimator is the Tobit model; but the Powell estimator leaves the distribution of ε_i unspecified:

$$y_i = \max\{0, x_i^{/}\beta + \varepsilon_i\} \qquad [5.2]$$

Following Powell (1986a), the first step in the definition of censored regression quantile estimators is to determine the functional form of the quantiles of y_i, i.e. $F_Y^{-1}(q|x_i, \beta)$. If the distribution of ε_i is continuously differentiable with positive density at the qth quantile of the i.i.d. sequence,

ε_i, i.e. $F^{-1}(q)$ and given the equivariance of population quantiles to monotonic transformations, the qth conditional quantile of y_i can be represented as:

$$F_Y^{-1}(q|x_i, \beta) = \max\{0, x_i^{/}\beta + F^{-1}(q)\} \quad [5.3]$$

where equation (5.3) is obtained by replacing ε_i in equation (5.2) with its qth conditional quantile, $F^{-1}(q)$. The censored quantile regression estimator of β_q, i.e. $\widehat{\beta_N}(q)$ is that value of β for each quantile (q), which minimizes the following loss function:

$$Q_n(\beta; q) = \frac{1}{N}\sum_{i=1}^{N} \rho_q\left[y_i - max\{0, x_i^{/}\beta\}\right] \quad [5.4]$$

where $\rho_q(x) = [q - 1(x \leq 0)].x$ are quantile-specific weights (Koenker & Bassett Jr, 1978).

Consistency of the qth regression quantile estimator of the slope coefficients requires that the distribution function of the error term be continuously differentiable with positive density at the qth quantile of the i.i.d. sequence, ε_i. Additionally, it must be that $x_i^{/}\beta + F^{-1}(q) > 0$ for a positive fraction of observations, while regressors satisfying this condition must have enough variation to identify β. Thus, Powell's estimator makes use of observations that are not likely to be censored (Kowalski, 2009).

In summary, the above Powell model uses semi-parametric conditional quantiles to characterize the distribution of the observable data in terms of the parameters and the distribution of the error term, as well as allowing censoring at zero earnings non-parametrically. This is in contrast with the fully parametric maximum likelihood estimation of the censoring process in the Tobit/Heckman model.

To also obtain average gradients in a censored framework, the following sample selection model was implemented following Cameron and Trivedi (2010).

Let latent participation (y_1^*) be a linear function of observable covariates (x_1) and a random unobservable term (ε_1) and let latent wage (y_2^*) be a linear function of observable covariates (x_2) and a random unobservable term (ε_2):

$$y_1^* = x_1^{/}\alpha + \varepsilon_1$$

$$y_2^* = x_2^{/}\beta + \varepsilon_2$$

The selection equation for observed participation (i.e. y_1) holds as follows:

$$y_1 = \begin{cases} 1 \; if \; y_1^* > 0 \\ 0 \; if \; y_1^* \leq 0 \end{cases} \qquad [5.5a]$$

while the outcome equation for observed wages (i.e. y_2) is:

$$y_2 = \begin{cases} y_2^* \; if \; y_1^* > 0 \\ - \; if \; y_1^* \leq 0 \end{cases} \qquad [5.5b]$$

i.e. actual wage is observed if the participation index exceeds zero and is unobserved otherwise (in reality, given that the outcome in the outcome equation is log wage, the censoring point is the smallest non-zero value, *K*, rather than zero). The model assumes normality of the errors as well as homoscedasticity.

5.4.2 Empirical model

Each of the above models was specified to include one health measure at a time. For the empirical specification of the censored quantile regression model, the determinants of wages are as follows: health (log of BMI (lnbmi), log of the CES-D10 psychological index (lncesd10) and SAH (sah) respectively for each equation), education, quadratic in age, location, race, gender, provincial unemployment rate, marital status, number of under-17 children in the household, tenure in current job, union membership, belonging to the managerial/professional occupational category, employment in the tertiary sector and number of weekly hours worked.

Given possible censoring of log wages, I used a cut-off, *K*, which is the smallest non-zero value of log wage. This is due to the fact that a censoring value of zero may not be feasible given that the outcome of interest is log wages. This is consistent with earlier work in the literature (Kowalski, 2009).

Corresponding Heckman selection models were also estimated to ascertain average relationships in a censored framework. Explanatory variables in the participation equation were: health (lnbmi, lncesd10 and sah respectively for each equation), education, quadratic in age, location, race, gender, provincial unemployment rate, marital status and number of under-17 children in the household. Furthermore, identifying restrictions involved the inclusion of household grant receipt and household size only in the participation equation as they are expected to directly influence the decision to work rather than actual wage received.

5.4.3 Sample, variable description and descriptive analysis

Only Africans and coloureds were included in the sample as the Indian and white samples were very small in most quartiles of the wage distribution. I restricted the sample to adults aged 20-56 years in wave 1 so as to exclude students and those likely to retire by wave 3 from the analysis as well as ensure comparability of results for both waves. Additionally, only employees were included in the wage regressions (i.e. the self-employed and casual workers were excluded). Wage, as used in this chapter, is monthly take-home pay from main job.

Along with other physiological health proxies, BMI (comprising weight and height measures) is regarded as an indicator of the extent of nutrient intake and absence of disease in childhood even though height has some genetic component (Husain, 2010). Similarly, Schultz (2002) has used height as an indicator of health human capital in a wage regression (the drawback of using height as an indicator of work capacity in a wage regression was discussed in the above empirical review). Therefore, I used BMI as a proxy for physical health. BMI has also been used in a study of the health-wage gradient in Ethiopia (Croppenstedt & Muller, 2000). Later in the study, I used an indicator of being underweight, following the WHO classification of various BMI ranges as shown below, to ascertain whether being underweight has a negative gradient with wages since such a condition partly presupposes lack of adequate body nutrients and/or health challenges.

The ten-question Center for Epidemiological Studies Depression Scale (i.e. c*esd10*) – a standard depression scale -, was obtained from a ten-question list eliciting an individual's psychological/emotional health. Responses to questions on whether respondents felt unusually bothered, had trouble keeping their minds on what they were doing, felt depressed, felt fearful, had restless sleep, felt everything was an effort, felt lonely, could not get going, felt happy and/or hopeful about the future (all in the previous week) were used to construct the c*esd10* variable. Each question had four options regarding the frequency of the experience in the past week, ranging from "rarely or none" (< 1 day) to "all of the time" (5-7 days). These questions are similar to those found in the British General Household Questionnaire, originally used to evaluate psychiatric illness (Contoyannis & Rice, 2001). Though the CES-D20 index (a twenty-question index) appears popular in the literature, the questions necessary for creating it are lacking in the dataset. But the CES-D10 index has been shown to be informative of one's psychological state. It ranges from 0 to 30, where a cut-off of 10 or more indicates depressive symptoms (Andresen, Malmgren, Carter, & Patrick, 1994; Zhang et al., 2012). Internal consistency of the c*esd10* index was achieved given a Cronbach alpha coefficient of 0.78 (0.74) in wave 1 (wave 3). Therefore, this variable was used as an indicator of psychological health. Robustness checks were later conducted with a CES-D8 index obtained by excluding the two questions on happiness and hope (Cronbach alpha coefficients of 0.84 and 0.90 in waves 1 and 3 respectively).

Data constraints prevented the use of longer term psychological health measures in the main analysis. I acknowledge that individuals who are generally free of psychological conditions but happened to frequently experience some of the above adverse conditions in the previous week are likely to be erroneously assigned a high CES-D10 measure, thus suggesting that they have depressive symptoms. This will likely result in an under-estimation of the gradient between psychological health/depression and wages, as their monthly wages are highly unlikely to be adversely affected by such rare conditions. Theoretically, this will result in the well-known attenuation bias (Cameron & Trivedi, 2005). I return to this issue in the results section where I show that a longer term measure of psychological health yields steep health-wage gradients.

SAH served as a measure of general health status as it evaluates overall health status as perceived by the respondent. This is informed by earlier literature which maintains that SAH is a summary/representative measure of overall health status (Benítez-Silva et al., 2004; Ferraro, 1980; LaRue et al., 1979; Nagi, 1969)

Kernel densities of the distribution of the natural logarithm of real wage in both waves (Figure 5.1) indicate that they are not exactly normal. This apparently suggests that the mean might not be a very good representation of the wage distributions, necessitating estimation at different parts of the wage distribution. As expected, males generally earned more than females in both waves, suggesting a male wage premium.

Figure 5.1: Kernel densities of natural logarithm of real wage in waves 1 and 3

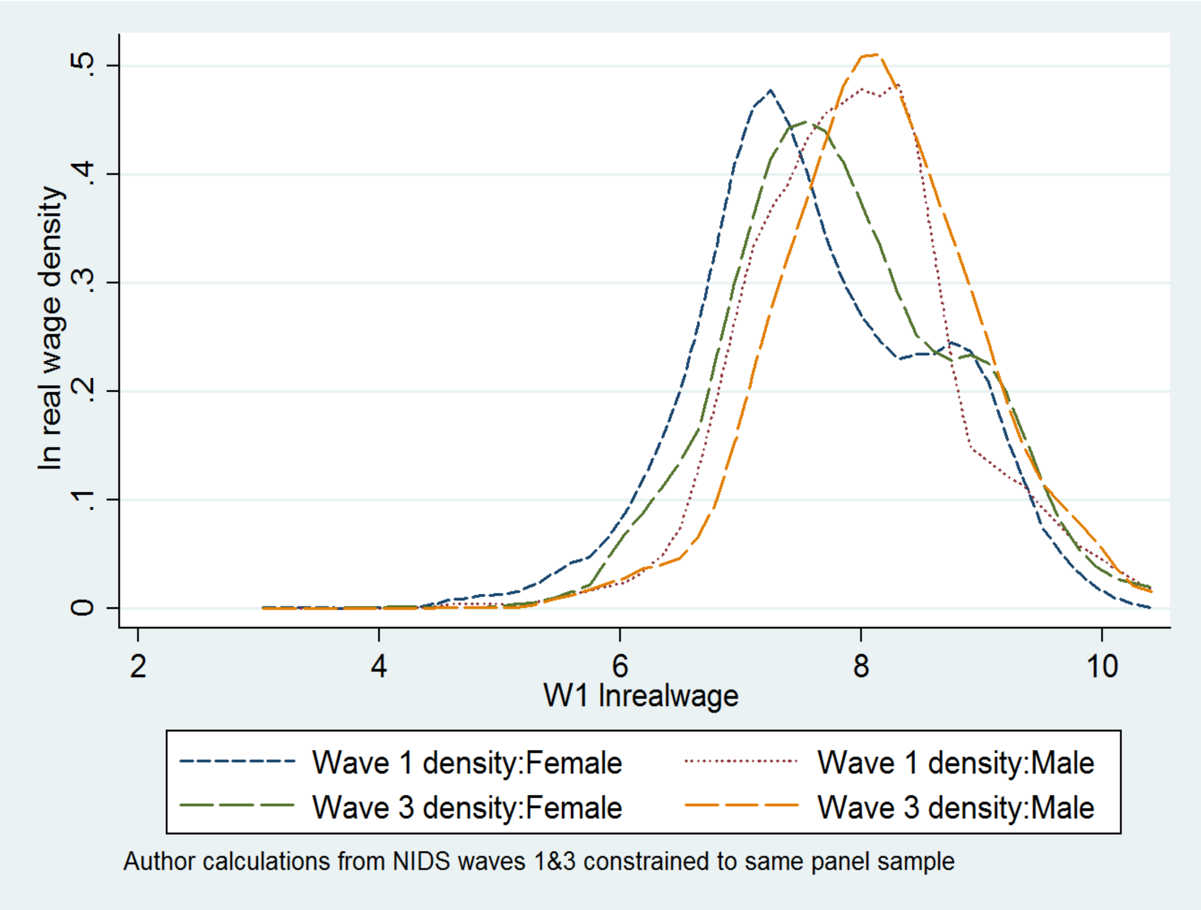


Table 5.1 depicts the descriptive statistics of variables used in the study.

Table 5.1: Descriptive statistics (working age Africans and coloureds)

	(1)	(2)	(3)	(4)	(5)	(6)
		Wave 1			Wave 3	
Variable	N	Mean	Std.Dev.[a]	N	Mean	Std.Dev.
employee†	7122	0.47	0.5	5437	0.58	0.5
sah	9443	0.83	0.4	7828	0.88	0.3
bmi	8448	26.07	6.4	7634	27.20	6.3
cesd10	9479	8.34	4.8	7824	7.13	4.4
real wage	3156	3993.00	6053.0	3054	4593.00	5353.0
matric	11321	0.31	0.5	8778	0.33	0.5
age	11400	34.29	10.0	8795	38.95	10.2
age^2	11400	1276.00	735.4	8795	1622.00	838.7
num. children††	11400	1.78	2.0	8795	1.45	1.7
prov. unemp‡	11400	23.58	3.0	8795	25.55	3.5
formal. loc‡‡	11400	0.57	0.5	8775	0.57	0.5
African	11400	0.89	0.3	8795	0.90	0.3
coloured	11400	0.11	0.3	8795	0.10	0.3
married	9481	0.43	0.5	7828	0.43	0.5
male	11400	0.48	0.5	8775	0.47	0.5
grant	11400	0.55	0.5	8775	0.57	0.5
household size	11400	4.70	3.3	8775	4.76	3.2
tenure	3088	6.87	7.7	3030	7.70	8.7
union	3076	0.34	0.5	3047	0.35	0.5
manprof¥	3115	0.19	0.4	3014	0.21	0.4
tertiary	2916	0.50	0.5	2919	0.63	0.5
hours	2880	40.58	14.8	3001	41.04	13.5

Source: NIDS wave 1and wave 3; author's calculations. Samples restricted to 20-56-year-old respondents in wave 1. Samples were corrected for survey design, national representativeness and non-random attrition; [a]standard deviation †dummy variable (=1 if respondent is an employee; 0 otherwise); ††number of under-17 children in the household; ‡provincial unemployment rate; ‡‡dummy variable (=1 if respondent resides in formal location; 0 otherwise); ¥dummy variable (=1 if respondent belongs to the managerial/professional occupational category; 0 otherwise)

As Table 5.1 indicates, average values/proportions of most variables did not change remarkably between both waves. There was a five percentage point increase in the proportion of the sample that self-reported excellent, very good or good health between waves 1 and 3. Also, average real wage increased by 15%. Average provincial unemployment rate worsened over time, a two percentage point increase. As noted in the Chapter 4, this variable is used as a proxy for local labour market conditions in an attempt to capture demand-side determinants of labour supply. Also, given the possible existence of (an often negative) wage curve depicting the relationship between regional unemployment and wages (Blanchflower &

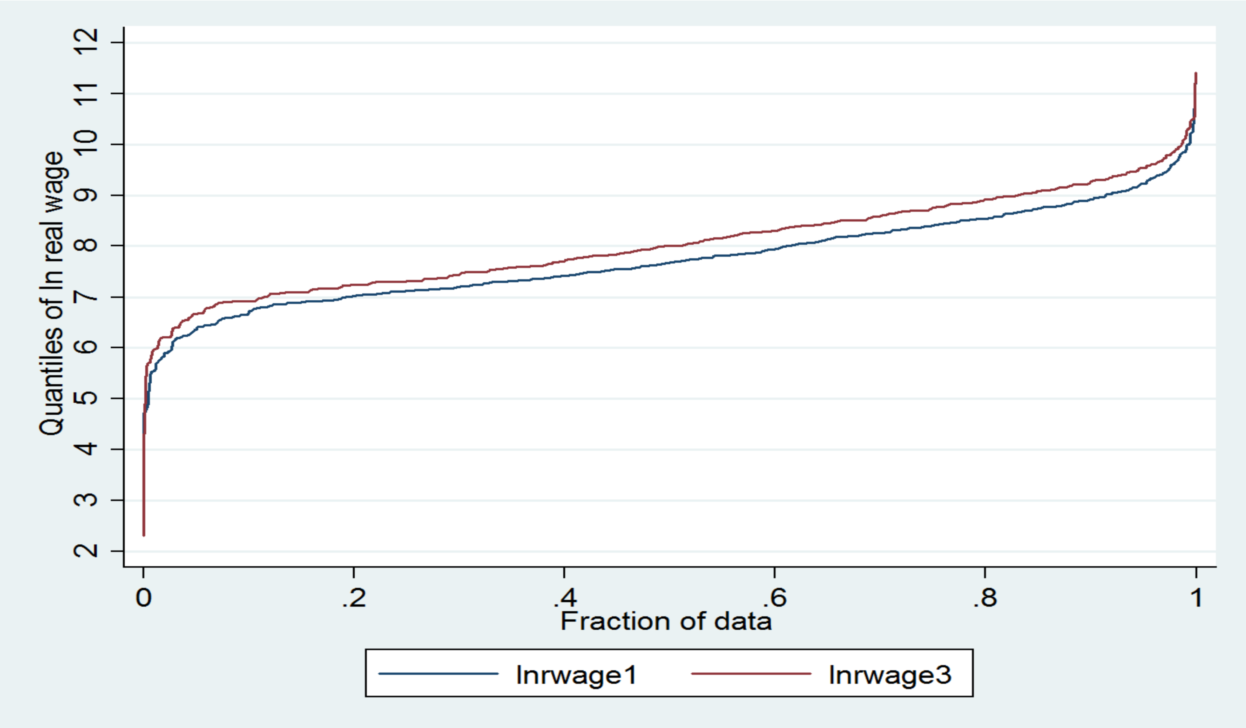
Oswald, 1994; Ilkkaracan & Selim, 2003), it is expected to be negatively associated with both employment and wages. While the racial composition of the sample remained virtually unchanged, the proportion of grant-receiving households increased by two percentage points. Though both racial groups analysed here were victims of racial segregation under apartheid, the huge disparities in the unemployment rates among them (Table 2.2 in Chapter 2) apparently suggests that race may be a significant determinant of employment. The same can be said about wages (Figure 2.5 in Chapter 2). As observed in Chapter 4, household grant receipt may negatively affect labour supply by imbuing beneficiary household members with higher reservation wage. Conversely, it may increase prime age labour supply (and hence employment) by providing resources for job search or providing the care-giving which might have initially kept prime age household members out of the labour market (Ardington et al., 2009). Therefore, its relationship with employment, especially in a South African context, is ambiguous a priori. Age is likely to be positively related to employment and wages. Similarly, residing in formal locations is expected to be positively related to employment and wages as informal locations have been historically associated with poor living standards in South Africa (Statistics South Africa, 2004). The proportion of the sample residing in formal locations remained constant at 57% across waves. Presence of young children is expected to be negatively associated with reduced labour supply (Dinkelman & Pirouz, 2002). More children (especially if they are biological) may increase wage through an increase in dependency allowance (which is part of take-home pay used as a measure of wage in this study). The average number of under-17 children in the household slightly declined from 1.8 to 1.5 between wave 1 and wave 3. The inclusion of marital status in the wage equation was informed by the literature on marriage premium especially for men (see e.g. Antonovics & Town, 2004). Tenure, union membership, belonging to the managerial/professional occupational category (relative to the semi-skilled and elementary classes), working more hours and employment in the tertiary sector (relative to the primary, secondary or private/exterritorial sector) are expected to be positively correlated with wages.

The sample became heavier on average, with average BMI increasing from 26.1 to 27.2. Given that the original BMI distribution was highly skewed, especially due to suspect responses (e.g. BMI values well above 100), the analysis was restricted to BMI values not exceeding 50 (BMI values above 50 constituted about 1% and 0.6% of non-missing BMI observations in the wave 1 and wave 3 datasets respectively). A log transformation (used in the regression analyses) was very close to the normal distribution (kernel density graph available on request). Additionally, such log transformation is necessary as a non-linear relationship is suspected between BMI and earnings given that not all BMI ranges are considered to have beneficial health consequences; for individuals aged at least 20 years, WHO classifies BMI values less than 18.5, 18.5-24.9, 25-29.9, and at least 30 as underweight, normal, overweight and obese respectively (Ardington & Case, 2009). Such a log transformation of BMI has also been used in a similar study in Ethiopia (Croppenstedt & Muller, 2000) and has the added advantage of yielding elasticities. A quadratic BMI specification may be argued to be suitable as it will enable the calculation of a turning point. But initial checks showed that such a quadratic specification did not suit the sample. This may not be surprising given generally non-decreasing association between BMI and income in developing countries as opposed to developed countries' experience (Wittenberg, 2011).

A look at the wage distribution suggests enough observations to carry out the proposed quantile analysis as the sample size in each wage quartile had approximately 790 and 765 observations in wave 1 and wave 3 respectively.

Figure 5.2 depicts quantile plots for waves 1 and 3 wage distributions.

Figure 5.2: Wage quantiles



Source: Author computation

Figure 5.2 (which can be viewed as an empirical CDF with reversed axes) depicts comparative wage distributions in waves 1 and 3 across different wage quantiles. This is consistent with, and more informative than, Table 5.1 (which only shows averages) as the wave 3 wage distribution dominated the wave 1 distribution.

Table 5.2- Table 5.4 compare average real wage between health (SAH), educational and racial groups respectively across wage quartiles.

Table 5.2: Healthy-sick average real wage differences across wage quartiles

	Wave 1			Wave 3		
		Average real wage (Rands)			Average real wage (Rands)	
Quartiles	N	Sick	Healthy	N	Sick	Healthy
1	786	676	752	768	924	921
2	782	1398	1399	758	1901	1876
3	784	2482	2648	767	3737	3448
4	782	6311	8929	760	8865	10339

Source: NIDS wave 1and wave 3; author's calculations. Sample is restricted to 20-56-year-old respondents in wave 1. Samples were corrected for survey design, national representativeness and non-random attrition

Table 5.2 reveals some important features. Firstly, average monthly real wage differential between the healthy and the sick across wage quartiles were not uniform in both waves. More formal tests of mean differences (not presented here) showed that though these differences were statistically significant at conventional levels, the strength of statistical significance declined with increasing quartiles in wave 1, as the third and fourth quartiles were barely significant at the 10% level while the bottom two quartiles were significant at 1% and 5% respectively. In wave 3, only the bottom quartile showed a statistically significant difference between the healthy and the sick. These findings apparently suggest some heterogeneity in the health-wage relationship across the wage distribution at least in absolute terms. Another feature of the data is that average real wage in each quartile increased temporally across waves, i.e. both the sick and healthy groups earned higher in real terms in wave 3 relative to wave 1 in each quartile. Even the sick in wave 3 earned more than the healthy in wave 1 except for the topmost quartile.

It is also important to examine descriptive evidence on the heterogeneity of the relationship between wage and key covariates. Table 5.3 conducts similar analysis as Table 5.2 for educational attainment (between respondents with and without at least a matric qualification).

Table 5.3: Matric vs. non-matric average real wage differences across quartiles

	Wave 1			Wave 3		
		Average real wage (Rands)			Average real wage (Rands)	
Quartiles	N	Non-matric	Matric	N	Non-matric	Matric
1	788	737	743	767	913	974
2	790	1384	1452	758	1854	1944
3	788	2556	2756	767	3386	3564
4	782	6071	9614	760	7430	11151

Source: NIDS wave1 and wave 3; author's calculations. Sample is restricted to 20-56-year-old respondents in wave 1. Samples were corrected for survey design, national representativeness and non-random attrition

From Table 5.3, wage differences between respondents with at least a matric and their less educated counterparts were statistically significant across all quartiles in both waves except for the lowest quartile in wave 1. Moreover, there was apparently a widening of education-related wage gaps over time in absolute terms. Also, though not shown in the table, the bottom three quartiles were mainly populated by those with low education levels while the converse obtained for the topmost quartile.

The above analysis was also conducted for racial groups in Table 5.4. In wave 1, significant wage differences occurred only at the two poorest quartiles, where Africans earned more than coloureds. The same was true in wave 3 (though coloureds significantly earned more than Africans in these lower quartiles).

Table 5.4: Race-based wage differences across wage quartiles

	Wave 1			Wave 3		
		Average real wage (R)			Average real wage (R)	
Quartiles	N	African	Coloured	N	African	Coloured
1	789	746	669	768	920	934
2	791	1402	1388	759	1869	1962
3	790	2637	2620	767	3462	3470
4	786	8584	9100	760	10088	11587

Source: NIDS wave1 and wave 3; author's calculations. Sample is restricted to 20-56-year-old respondents in wave 1. Samples were corrected for survey design, national representativeness and non-random attrition

Finally, Table 5.5 below depicts the means/proportions of the different health measures analysed in this study, across wage quartiles. The purpose is to tentatively ascertain whether any relationship is likely to exist between each of them and wage as well as the likely direction of such relationship. From the table, it is apparent that a positive relationship likely exists between SAH and wages, while a negative relationship is expected between *cesd10* and wages. These findings clearly follow a priori expectations. There was no clear-cut pattern for BMI, though the highest earner group had the highest BMI value in each wave. High BMI among top earners may not be surprising in a developing country context given observed positive relationship between BMI and wellbeing in developing countries (Wittenberg, 2011).

In summary, the foregoing analysis gives tentative indication of differing strengths of association between general health status (proxied by SAH) and wages over the wage distribution. This makes it even more imperative to analyse whether an analysis which focuses on different points of the wage distribution may reveal important features of this relationship which are likely to be masked in a mean-based analysis. Additionally, analysis at different points of the wage distribution will enable one to ascertain whether controls affect wages differently at different points across the wage distribution as suggested by the foregoing descriptive analysis.

Table 5.5: Distribution of health indicators across wage quartiles

	BMI				CESD10				SAH			
	Wave 1		Wave 3		Wave 1		Wave 3		Wave 1		Wave 3	
Quartiles	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N
1	25.8	705	26.9	758	8.8	788	7.0	768	0.81	786	0.89	768
2	25.5	685	26.7	738	7.6	788	7.0	759	0.90	782	0.92	758
3	25.6	677	26.4	744	7.2	787	6.1	767	0.91	784	0.95	767
4	27.6	653	29.3	738	6.5	784	5.9	760	0.91	782	0.95	760
Total	26.2	2720	27.4	2978	7.4	3147	6.4	3054	0.89	3134	0.93	3053

Statistics corrected for national representativeness and non-random attrition

The approach adopted in the following sections was to estimate OLS regressions of wage on each of the health measures and relevant controls. Thereafter, specifications accounting for censoring were implemented in both mean- and quantile-based analyses, where the latter examined the relationships among the first quartile, median and top quartile of the wage distribution.

5.5 ECONOMETRIC RESULTS AND DISCUSSION

5.5.1 Ordinary least squares estimates

Table A5.1 in the appendix depicts OLS regression results for both waves. The results suggest positive association between physical health (proxied by BMI) and general health condition (proxied by SAH) on the one hand, and wages on the other, as well as negative gradients between adverse psychological health and wages. As shown in the table, coefficients of health measures were stable across waves while coefficients of regression controls were robust to both health measure and wave.

5.5.2 Accounting for censoring

Earlier, it was noted that the nature of wages makes it susceptible to censoring given that one cannot observe wages for the unemployed. If censoring exists, the sub-sample of wage earners will no longer be a random sample of labour market participants and failure to account for such censoring will lead to inconsistent estimates (Cameron & Trivedi, 2005). Given that 56% (44%) of the sample was left-censored in wave 1 (wave 3), I present results from Heckman sample selection models (capturing the average gradient) and censored quantile regression models (allowing for heterogeneity of association) for each of the health measures in both waves. This enables a comparison of the results across time, model and health measure.

Heckman sample selection results for waves 1 and 3 are depicted in Table A5.2 and Table A5.3 in the appendix. As earlier indicated, identification was achieved by excluding grant receipt (i.e. non-labour income) and household size from the wage equation (non-labour income has also been previously used to identify participation in a similar context (Jones et al., 2006)). Both tables show evidence of negative sample selection bias if censoring was not controlled for, i.e. if the possibility that the sub-sample of wage earners may be different from the unemployed was ignored. Ordinarily, this goes against conventional wisdom as it suggests that unobserved determinants of employment are negatively correlated with unobserved determinants of employee earnings. However, a possible explanation lies in South Africa's high graduate unemployment rate; the result of a structural mismatch between available jobs and the skill set of the unemployed (African Economic Outlook, 2012). Therefore, there appears to be a glut at menial low-paying jobs often performed by low ability individuals, and acute labour scarcity at jobs requiring high end skills even though many graduates remain unemployed. Consequently, it appears that low ability individuals have a higher probability of being employed than their high ability (and more educated) counterparts.

A 10% BMI increase was associated with 2.1-2.6% wage increase. Also, a 10% increase in the CES-D10 index was associated with a 0.5% decrease in wages (only in wave 1), while self-reporting good, very good or excellent health was associated with 11-15% wage rise relative to self-reporting fair or poor health.

To obtain more nuanced estimates of the various health-wage gradients, censored quantile regression models were estimated - Table 5.6 - (full results are reported in Table A5.4 and Table A5.5 in the appendix).

Table 5.6: Censored quantile regression results of health-wage gradient across health definitions

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8	(9)
Q25	Q50	Q75	Q25	Q50	Q75	Q25	Q50	Q75
	lnbmi			lncesd10			sah	
				Wave 1				
	N=2189			N=2518			N=2505	
0.31***	0.33***	0.37***	-0.07*	-0.06*	-0.06	0.22**	0.17***	0.20***
(0.12)	(0.08)	(0.11)	(0.03)	(0.03)	(0.05)	(0.07)	(0.07)	(0.06)
				Wave 3				
	N=2747			N=2806			N=2805	
0.33***	0.41***	0.45***	-0.06*	-0.07**	-0.05	0.22*	0.17**	0.22***
(0.11)	(0.09)	(0.10)	(0.04)	(0.03)	(0.03)	(0.12)	(0.09)	(0.07)

Cluster-robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1

Table 5.6 shows statistically significant gradients in both waves and across all quantiles for the different health measures. All coefficient signs also conformed to a priori expectations. A 10% increase in BMI was associated with 3.1%-3.7% (3.3%-4.5%) wage increase in wave 1 (wave 3). These results are not very different from findings in Europe, where a 10% BMI increase was associated with real earnings reduction of 3.3% and 1.9% for males and females respectively (Brunello & d'Hombres, 2007). They are however smaller than that found among Ethiopian farmers, where a 10% BMI increase was associated with 27% wage increase (Croppenstedt & Muller, 2000). This difference possibly emanates from differences in context as Croppendstedt and Muller examined the gradient for farmers (for whom physical strength is arguably a very important requirement for work) while this study focuses on all wage employees. To further buttress these differences, only 8.9% (7.2%) of respondents in this study were involved in the agricultural sector in wave 1 (wave 3); and that, even in an apparently more mechanized agricultural sector compared to what obtains in Ethiopia. In another study, Schultz (2002) found that a one centimetre height increase was associated with 1.5% (1.7%) wage increase for men (women) in Ghana. Though these estimates may not be directly comparable as noted in the empirical review above, these indicate that anthropometric measures of health are positively associated with higher earnings in Africa.

Similarly, a 10% increase in the CES-D10 index was associated with a 0.6%-0.7% (0.7%) decline in real wages across the quantiles in wave 1 (wave 3). These CES-D10 coefficients are similar to the above mean-based estimates. As noted in the empirical review, serious mental illness was associated with annual earnings reduction of about USD16300 in the US (Kessler et al., 2008) while psychological conditions in childhood permanently lowered earnings by about USD4094 per year in the US (Smith & Smith, 2010). Also, serious mental illness sufferers earned a third less than median earnings in a survey of 19 countries including some in Africa (Levinson et al., 2010). Though direct comparison of these estimates with the ones in this study is not intended due to fundamental differences in study design including psychological health measurement, these results validate my finding that psychological health is an important determinant of wages. Also similar to the findings in this study, mental health did not appear to significantly affect earnings among top earners for a number of mental health conditions like anxiety disorder and anti-social personality disorders in the US (Marcotte & Wilcox-Gok, 2003). The authors explained significant mental health effect at lower quantiles and insignificance at higher quantiles as possibly emanating from limited access to mental health services for low-paid workers as well as better working conditions (e.g. more paid sick leave) for salaried high earners relative to their low-earning hourly paid counterparts. Unfortunately, there is hardly any developing country empirical evidence on the mental health-wage gradient to the best of my knowledge. Regressions using non-logged CES-D10 measures also conformed to a priori expectations while statistical significance at conventional levels were only attained at the median as in Table 5.6 (available on request).

Also, self-reporting excellent, very good or good health was associated with 17-22% wage increase relative to reporting fair or poor health in both waves. There was no consistently monotonic pattern in the coefficient size across the wage distribution for psychological and overall health variables. However, coefficient size increased monotonically for BMI in both waves, a result at variance with Rivera and Currais (2005) who noticed decreasing gradients as one moved up the wage distribution in Brazil. This might be due to the health measure used by Rivera and Currais, an index of health limitation, which is likely to affect low-earning workers (who are often engaged in physically demanding tasks) more than high earners.

In comparing the magnitudes of the gradient between wages on the one hand and SAH and psychological health on the other, Contoyannis and Rice (2001) found that the SAH-wage gradient exceeded that of psychological health (though the former was only statistically significant for females). Though both gradients are not directly comparable in this study given that SAH is a dummy variable while psychological health is continuous, the tenuous statistical significance as well as insignificance in some quantiles for the latter makes one inclined to accept the existence of steeper general health gradient relative to psychological health. Also, Table 5.6 indicates nontrivial numerical heterogeneity for BMI especially in wave 3, while there was not much heterogeneity across the quantiles for general health and psychological conditions. The above results show that mean-based coefficient estimates were lower than quantile-based estimates. These results were generally replicated when all three health measures were included in the same regression except that *lncesd10* was no longer statistically significant at conventional levels in both waves. These latter results are available on request.

Given that the BMI and CES-D10 coefficients measure elasticities, the foregoing shows that the gradient between physical health and wages exceeded that of psychological health. This is not surprising when considered in light of the above point that the cross sectional psychological health variable likely suffered from nontrivial measurement error (which will result in the attenuation of the wage-*cesd10* gradient).

5.5.3 Robustness checks

To obtain arguably more exogenous (i.e. less error-ridden) estimates, I constructed a dummy variable measure of long term psychological/depression condition, *ltdep*. Given Andresen et al.'s (1994) suggestion that a CES-D10 measure of 10 or more indicates depressive symptoms, I considered an individual as suffering from long term psychological condition/depression if her CES-D10 index was at least 10 in both waves. I excluded those with CES-D10 measure in only one wave. This measure is likely to be less error-ridden and thus represent true psychological health condition more accurately than the cross sectional measure. This is because, individuals with such depressive scores over both waves are likely to be truly suffering from some psychological condition. This resulted in a sample size of 2536 of which 7.4% (190 respondents) had long term psychological/depressive condition(s). The resulting censored quantile regression results are depicted in the first three columns of Table 5.7.

Similarly, to examine the long term gradient between SAH and wages, I constructed an index of long term health, where respondents who self-reported fair or poor health (the so-called sick) in both waves were classified as being long term sick (*ltsick*). On the other hand, those who reported excellent, very good or good health (the so-called healthy) in both waves were deemed to not being long term sick. This resulted in a sample size of 3794 of which 3.6% (138 respondents) reported being "long term sick". The results are presented in columns 4-6 of Table 5.7.

Table 5.7: Censored quantile regression results of gradient between long term health conditions and wages

	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
	Q25	Q50	Q75	Q25	Q50	Q75
		ltdep† (N=2330)			*ltsick*†† (N=1950)	
	-0.15	-0.23***	-0.21**	-0.32	-0.27**	-0.21
	(0.10)	(0.07)	(0.08)	(0.25)	(0.13)	(0.21)
Pseudo R^2	0.26	0.34	0.37	0.27	0.34	0.37

Cluster-robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1; † dummy variable (=1 if the respondent is long term depressed; 0 otherwise); †† dummy variable (=1 if the respondent is long term sick; 0 otherwise)

Table 5.7 shows that the signs of all *ltdep* and *ltsick* coefficients conformed to a priori expectations. Though coefficients of *ltdep* and *lncesd10* cannot be directly compared given the categorical and continuous nature of both variables respectively, the coefficients of *ltdep* exhibited stronger statistical significance compared to *lncesd10*. Also, *ltdep* coefficients were similar to their SAH counterparts. Individuals with depressive symptoms in both time periods suffered wage penalties of about 21-23% relative to their healthier counterparts. Longer term general health (*ltsick)* coefficients were mostly larger than their cross-sectional counterparts though they were largely statistically insignificant. These latter coefficients (*ltsick)* should however be viewed with caution given very low variation (as indicated above, the "long term sick" constituted only 3.6% of the sample). This low variation resulted in higher standard errors (relative to cross-sectional estimates) thereby rendering all but the median coefficient statistically insignificant. Similar regressions for experiencing long term underweight could not be estimated due to very low variation, as only 1.7% of the sample were consistently underweight over the two waves.

On whether gradients persisted over time, I estimated wave 3 censored quantile wage regressions as a function of wave 1 health variables and wave 3 controls. This exercise is especially important given the earlier observation that BMI may be determined by socio-economic status indicators like wealth (which can be enhanced by wage income) and employment especially among blacks in South Africa (Wittenberg, 2011). A similar point applies to the other health measures. Such analysis will likely yield more exogenous gradients given that the health condition in question preceded wages by four years. The results, shown in Table 5.8, below indicate persistent gradients for physical health across the different estimated quantiles, though attenuated relative to the cross-sectional gradients as expected. Similarly, psychological health gradients were statistically significant at conventional levels across all estimated quantiles except the bottom quartile, while only the median gradient was significant at 10% for general health status. Apparently, these results suggest that the earlier observed gradients were not mainly driven by simultaneity between wage and health, especially for BMI and *cesd10*. A 10% increase in BMI in wave 1 was associated with 2.1-2.4% wage increase in wave 3 while such change in *cesd10* was associated with a 0.8% wage rise.

Table 5.8: Censored quantile regressions of current (wave 3) wage on past (wave 1) health measures

	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
	Q25	Q50	Q75	Q25	Q50	Q75	Q25	Q50	Q75
	Past lnbmi (N=2054)			Past lncesd10 (N=2330)			Past sah (N=2321)		
	0.21**	0.24**	0.24**	-0.07	-0.08**	-0.08**	0.10	0.12*	0.01
	(0.10)	(0.10)	(0.12)	(0.04)	(0.04)	(0.04)	(0.08)	(0.07)	(0.09)
Pseudo R^2	0.27	0.34	0.37	0.26	0.34	0.37	0.26	0.34	0.37

Cluster-robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1

Apart from using the log of BMI, another way to exploit possible non-linearities in the wage-body mass gradient is to estimate the gradient between wages and a discrete variable denoting being underweight (i.e. *underweight*). Given the earlier observation by Husain (2010) that BMI is an indicator of the extent of nutrient intake, it is expected that being underweight is associated with reduced productivity, hence lower wages, therefore resulting in a negative gradient. To empirically explore this hypothesis, I estimated an OLS regression of wage as a function of *underweight* and all controls included in the main regressions[10]. The results indicate that being underweight was associated with 22% (30%) wage decline (p<0.01) relative to being heavier in wave 1 (wave 3). This is

[10] Result is available on request. Though I also estimated censored quantile regressions of this relationship, the results are similar to the OLS results in terms of sign and statistical significance especially at the top quartile but these are not reported given very small sample of the underweight, especially in wave 3 where there were only 96 underweight wage earners in the sample.

consistent with the above positive and significant gradient between higher body mass and wages.

I also used the CES-D8 depression index to test the robustness of the psychological health measure. OLS wage gradients in wave 1 and wave 3 were -0.04 and -0.06, significant at 5% and 1% respectively. For the Heckman selection model, only the wave 3 gradient (-0.03) was significant ($p<0.1$). The same is true of the censored quantile model, as none of the wave 1 gradients was statistically significant at conventional levels. In wave 3 however, the lower quartile and median gradients were -0.07 (significant at 5% and 1% respectively). This is similar to the wave 3 median gradient for the CES-D10 index.

A gendered analysis accounting for censoring (see Table A5.6 in the appendix) suggests more nuanced relationships than are obvious in the foregoing analysis. It shows that the BMI-wage gradients were driven by males as there was hardly any statistically significant relationship among females in both waves. This apparently indicates significant returns for heft among males only. This is not surprising as men are generally employed in more physically demanding jobs relative to women, and therefore are likely to suffer higher wage penalties to physical health deficiencies. There was no such stark pattern for psychological health even as most coefficients were insignificant or barely significant statistically. Finally, gradients for overall health status were generally positive and statistically significant across the entire wage distribution. In Britain however, Contoyannis and Rice (2001) found significant association between psychological health and hourly wages for only males.

From the foregoing and considering the main model (i.e. the censored quantile model), one can deduce that the BMI-wage gradient was quite robust to model and time period as it retained its positive sign and statistical significance across quantiles. For SAH, coefficients were remarkably stable between waves 1 and 3. The same is true for *lnced10* though they were largely barely significant at conventional levels. Using longer term measures, psychological health gradients were shown to be about as steep as the cross-sectional general health gradient while longer term general health gradients largely lost statistical significance due to minimal variation (the magnitudes were mainly larger than their cross-sectional counterparts though). Using arguably more exogenous past physical and psychological health measures (relative to their cross-sectional counterparts), gradients generally remained strong and statistically significant while displaying minimal heterogeneity. Mean-based gradients were also generally stable for the different health measures over the four-year period but attenuated relative to the more informative quantile estimates.

For the regression controls (refer to the censored quantile regressions in Table A5.4 and Table A5.5), their signs largely conformed to a priori expectations while magnitudes were generally stable across waves. Magnitudes were also generally robust to health measure. A one year increase in job tenure was associated with 1-2% wage increase across quantiles. This is same as found in the UK (Jones et al., 2006). Also, having at least a matric qualification was associated with 54-61% wage increase in the two waves. Prior evidence indicated that education premium for African males and females with post-secondary education was 29% and 37% respectively, while that of coloured were 19% and 31% respectively in 1993 (Mwabu & Schultz, 2000). Managers/professionals had nontrivial wage premium over those in other occupational categories especially in wave 3, earning up to 46-53% more. In general, respondents in formal locations significantly earned higher than informal dwellers; formal location dwellers earned 9-22% more than informal dwellers. In a similar study, rural dwellers were found to earn 36-42% (Africans) and 11-100% (coloureds) less than their non-rural counterparts in South Africa in 1993 (Mwabu & Schultz, 2000). Though these cannot be directly compared with estimates in this study given differences in the two measures as well as the fact that Mwabu and Schultz estimated separate regressions for Africans and coloureds, it is however obvious that the role of location in wage determination appears to have narrowed over time in South Africa, perhaps indicating some spatial convergence. Such narrowing notwithstanding, the returns to formal location dwelling has important welfare implications given the apartheid legacy whereby most jobs were located away from informal areas, necessitating substantial transport costs for informal location dwellers. High commuting costs, together with lower wages will likely result in substantial welfare loss for such individuals and their families. The statistical insignificance of race may be due to the fact that the two racial groups considered in this analysis have been victims of racial segregation in the past. The gender gap was substantial across quantile and wave, as being a man was associated with 23-36% (25-41%) wage increase relative to being a woman in wave 1 (wave 3). A consistent pattern over the two waves was that returns to males was highest among the topmost quartile irrespective of the health measure being considered. Respondents that were married/living with a partner generally earned more than their "single" counterparts at the topmost quartile in wave 3. This suggests even more welfare gains to marriage/cohabitation given possible economies of scale (assuming such couples do not have more dependants than the "single" on the average). However, this relationship was not consistently statistically significant. The higher union premium uncovered in this study (43-51% across quantiles and health measure in wave 3) relative to previous evidence (for instance, Butcher and Rouse (2001) found a 20% union wage premium among African workers) might not be unconnected to evidence regarding unions persistently obtaining wage increases for their members in South Africa (Banerjee et al., 2008). Finally, more children in the household was associated with lower wages.

5.6 CONCLUSION

This chapter has established positive and statistically significant gradients between better health (physical, psychological and general) and wages for Africans and coloureds in South Africa even after controlling for education and other important wage determinants like occupational category, industry, union membership and gender. These gradients range from an elasticity of -0.06 to -0.07 for psychological health/depression to 0.31-0.45 for BMI in the short term. Persistently adverse general health and psychological conditions exhibit steep gradients though those associated with general health status are statistically insignificant at conventional levels save for the median. The gradients do not generally exhibit much heterogeneity save for physical health in 2012 (i.e. wave 3). Also, mean-based estimates are generally smaller than the more informative quantile estimates. Furthermore, the physical and psychological health-wage gradients are persistent across the estimated quantiles even after four years, as previous health conditions significantly predict current wages, while that of general health is only significant at the median. Gender-based analysis show steeper physical health-wage gradients for males relative to females. Thus, the study shows the existence of nontrivial relationship between physical, psychological and general health conditions on the one hand, and wages on the other, in South Africa. Also, there remain substantial wage premiums for being a professional/manager relative to being semi-skilled or performing elementary tasks, coupled with nontrivial union and male wage premiums.

CHAPTER 6

6 IMPAIRMENT-RELATED WAGE DISCRIMINATION IN THE SOUTH AFRICAN LABOUR MARKET

6.1 INTRODUCTION AND BACKGROUND

The last two chapters have established positive relationships between health on the one hand, and the willingness to participate in the labour market as well as wages earned upon securing wage employment, on the other. With regard to wages, the theoretical review in Chapter 3 noted that differences in human capital play an important role in determining productivity/wage differentials. And given that health constitutes part of human capital, it stands to reason that health should play a role in determining wage differentials. However, it was also pointed out in chapter 3 that wages may differ for non-productive reasons, one of which is discrimination (defined in this context as a situation where identical individuals in similar jobs are paid differently because of non-productive characteristics (Altonji & Blank, 1999)). It was mentioned that some of the markers of wage discrimination are gender, ethnicity and disability/impairment.

The need to combat health-related discrimination has led to the enactment of legislations like the Americans with Disabilities Act, 1990 as well as the UK Disability Discrimination Act, 1995 (Acemoglu & Angrist, 2001; Madden, 2004). Having recognized that apartheid institutionalized discrimination in many facets of national life especially on the basis of race, post-apartheid South Africa has also passed a number of anti-discrimination laws. These include the Employment Equity Act, 1998 (and its various amendments) and the Promotion of Equality and Prevention of Unfair Discrimination Act, 2000. While the former was mainly concerned with addressing employment-related discrimination, the latter was a more comprehensive legislation which addressed issues regarding discrimination by both public and private organizations and individuals on the basis of race, gender, disability, marital status, religion, etc. Therefore, issues of discrimination in arguably most aspects of national life have been taken care of from a policy perspective.

The foregoing policy framework notwithstanding, there is need to empirically ascertain whether discrimination on the basis of key characteristics and attributes still exist in the country. An important form of discrimination is wage discrimination and empirical evidence on the issue is mixed (Baldwin & Johnson, 1994; DeLeire, 2001; Jones et al., 2006). I examine such literature in the literature review section of this chapter and subsequently ascertain the existence or otherwise of unexplained impairment-related wage gaps (often dubbed wage discrimination) in South Africa.

6.2 RESEARCH AIM AND OBJECTIVES

The aim of this chapter is to ascertain whether impairment-related wage discrimination exists in South Africa. Specifically, the study intends to uncover:

i) The magnitude of average impairment-related differences in returns to characteristics (loosely termed wage discrimination) in South Africa.
ii) Whether such discrimination changed between 2008 and 2012.
iii) Whether such discrimination was heterogeneous across the wage distribution.

6.3 CONTRIBUTIONS OF THE STUDY

Studies on the existence of wage gaps/wage discrimination in South Africa have focused on the effect of gender, race and union membership in engendering such gaps/discrimination (Mwabu & Schultz, 2000; Schultz & Mwabu, 1998b). This, to the best of my knowledge, is the first study to investigate the existence of impairment-related wage gaps and discrimination in South Africa. Indeed, in ascertaining the wage premium associated with education, region, race and sex in South Africa, Mwabu and Schultz (2000) highlighted the need to investigate the health-related premium as well. Therefore, this study fills this research gap. From a policy perspective, the results of the study will enhance knowledge of the obviously complex nature of discrimination in South Africa and improve policy makers' ability to tackle the problem of inequity, a major goal of the South African government.

6.4 EMPIRICAL REVIEW OF THE HEALTH-RELATED WAGE DISCRIMINATION LITERATURE

This section focuses on a review of empirical evidence regarding health-related wage discrimination. This is to keep the discussion focused given the vastness of the empirical wage discrimination literature. Where non-health related literature is reviewed, it will be based on methodological grounds (i.e. if it has direct implications for the methods to be applied in this study). As the following review indicates, the workhorse model of wage discrimination studies in economics is the Blinder-Oaxaca model (Blinder, 1973; Oaxaca, 1973).

Johnson and Lambrinos (1985) and Baldwin and Johnson (1994) found, using Blinder-Oaxaca decompositions that handicapped men (women) in the US received a wage that was 44.5% (85%) that of their non-handicapped counterparts on the average, while a third of these differentials were unexplained by included characteristics. "Handicap" was defined to include individuals who suffered from diseases that have been demonstrated as evoking prejudice in attitudinal studies, like blindness or deafness. Also using the Blinder-Oaxaca decomposition on the 1984, 1992 and 1993 panels of the Survey of Income and Program Participation in the US, DeLeire (2001) found that among health-impaired men who self-reported that their impairment did not affect their work, only 3.7 percentage points of the earnings gap was explained by discrimination. Also, the amount of discrimination did not decrease between 1984 and 1993. Maranto and Stenoien (2000) have noted that obese individuals are not as well protected against labour market discrimination as people suffering from other forms of discrimination, under the Americans with Disabilities Act. Noting that men did not experience wage penalties until their weight exceeded standard weight by 100lb (where standard weight was defined as a BMI which equalled 21.1 or weight equal to 127 lb for a 5ft 5in woman, and BMI equal to 21.7 or weight equal to 162 lb for a 6ft man), they maintained that mildly obese white women in the US experienced greater wage penalties than black men experienced for weight exceeding 100% of standard weight. A drawback of this study however was the use of ordinary least squares log wage regression which was clearly inadequate to uncover wage discrimination.

A shortcoming of many studies on disability-related wage discrimination is the failure to account for possible non-random selection into wage employment when decomposing wage gaps. To fill this gap, Kidd et al. (2000) estimated wage gaps based on selectivity-corrected wage regressions. The results showed the existence of substantial wage and labour force participation differences between able-bodied and disabled men in the UK. Madden (2004) accounted for selection into both health and labour market status in the UK. The result suggested that accounting for selection into health status was of little empirical importance, while accounting for selection into labour market status and the direct impact of health on productivity resulted in a fall in measured discrimination. Using selectivity-corrected wage gap decompositions, O'Hara (2004) found evidence of double discrimination against American women as they were subject to gender- and impairment-related wage discrimination in the early stages of employment. Jones et al. (2006) found evidence of impairment-related wage discrimination in the UK especially among women. Exploiting differences in the work-limiting nature of a disability, Jones (2006) found that unobserved productivity differences played an important role in engendering employment gap in the UK. Though these studies corrected for sample selection, a drawback of most of them is that they netted out the selection term from the raw wage gap, thereby obtaining only the offered wage gap rather than the actual/observed wage gap, as in Reimers (1983).

While maintaining that sample selection should be accounted for when performing wage decompositions, Neuman and Oaxaca (2003) averred that netting out the selection term from a typical raw wage gap is not ideal as such a method would only yield selection-corrected average wage gap and not the more policy-relevant average observed wage gap (more on this in the theoretical model specification section). Using data from Israel, Neuman and Oaxaca (2003) found larger gender-related wage differentials than ethnic-based wage differentials. They suggested that the difference between average wage gaps when controlling for selection relative to when it was not controlled for could be quite high. Also, the selectivity-corrected gender- and ethnic-related wage gaps were found to be substantial.

Apart from non-correction for sample selection, another shortcoming of traditional Blinder-Oaxaca decomposition is that it only focuses on the mean wage gap, ignoring what happens at other points of the wage distribution. Juhn, Murphy and Pierce (1993) have proposed an estimation technique to correct for this. Using this procedure, Beblo et al. (2003) found that endowment effects became more prominent at higher deciles of the distribution of the gender-based wage gap in Germany, while it varied greatly over the wage distribution in France.

But, given that the model based on Juhn et al. (1993) is not robust to heteroscedasticity, a semi-parametric heteroscedasticity-robust decomposition technique (based on quantile regression) that decomposes unconditional quantiles has been adopted in decomposing wage gaps across time and between gender groups (Antonczyk, Fitzenberger, & Sommerfeld, 2010; Melly, 2005). Using this technique to decompose changes in the wage distribution between 1973 and 1989 in the US, Melly (2005) found that a nontrivial proportion of the increase in inequality over this period was due to differences in returns to skills, especially education, while residuals accounted for 20% of the increase in inequality in the 1980s.

To the best of my knowledge, empirical evidence on impairment-related wage gaps/discrimination in developing countries is limited; most studies focus on disability-related employment gaps/discrimination. Mitra and Sambamoorthi (2008) found no evidence of disability-related wage discrimination among men in rural Tamil Nadu, India. Rather, there existed disability-related employment discrimination. Also in a study of discrimination against the disabled in fifteen developing countries, Mizunoya and Mitra (2013) found that in those countries characterized by a disability-related employment gap (i.e. in nine countries), observable characteristics did not explain most of such gaps. Most developing country studies found lower employment rates among the disabled relative to the non-disabled (Hoogeveen, 2005; Palmer et al., 2012; Trani & Loeb, 2012).

A number of studies have also been conducted on wage gaps in South Africa. On wage discrimination and inequality, Schultz and Mwabu (1998b) have noted that the average wage between white and black workers differed by a factor of five, though half of this differential could be attributed to educational and location differences across the races (see also Mwabu & Schultz, 2000)). Also, union membership was associated with higher wages at the lower end of the wage distribution. Furthermore, detailed decompositions conducted by Leibbrandt and Woolard (2001) showed that household inequality in South Africa was closely linked to wage differences. A similar conclusion was also reached by Leibbrandt et al. (2011). Grun (2004) found evidence of employment discrimination against African women, while white women were likely to suffer from wage discrimination in the post-apartheid period. Furthermore, there is evidence of decline in the rate of employment of people with disabilities in South Africa; such decline has been mainly attributed to increase in disability grant pay-outs (Mitra, 2008). However, in spite of such substantial evidence of labour market-related inequality, there is hardly evidence of a possible association between impairment and wage gaps in South Africa.

6.5 ECONOMETRIC ANALYSIS

6.5.1 Theoretical model specification

The first model estimated was a selectivity-corrected Blinder-Oaxaca decomposition model of average (log) wage differences as in Neuman and Oaxaca (2004). This is due to plausible sample selection bias arising from non-random selection into wage employment and was intended to capture the average wage gap. Furthermore, differences in the unconditional wage distribution based on quantile regression model a la Melly (2005) were decomposed.

A major challenge in conducting studies of health-related wage gaps/discrimination is the possible endogeneity of health. Unlike race and gender (some of the popular measures on which discrimination is analysed) which can be plausibly argued to be exogenous, some health conditions are the result of choices. Moreover, health can influence wages just as wages and working conditions can affect health. These relationships can theoretically result in health endogeneity in a wage regression. However, Madden (2004), who carried out similar analysis in the UK noted that accounting for health endogeneity made little difference relative to when endogeneity was not controlled for. Again, Ettner (2000), who adopted a two-step IV model on US data found that the effects of health on labour market outcomes were not sensitive to reverse causation (see also Jones et al., 2006). Therefore, given the lack of suitable instruments in the dataset as well as the foregoing findings, this analysis did not account for the possible endogeneity of impairment to wages.

For the selection-corrected Blinder-Oaxaca decomposition, the following model of employment and wage decisions was specified:

$$emp_{ij}^{*} = G_{ij}^{/}\beta_j + \varepsilon_{ij} \qquad [6.1]$$

$$w_{ij} = X_{ij}^{/}\alpha_j + \vartheta_{ij} \qquad [6.2]$$

where emp_{ij}^{*} and w_{ij} are respectively a latent variable associated with employment and wage for individual i in impairment status $j = \{NI, I\}$; G and X are predictors of employment and wages respectively for individual i in health status *j*. β_j and α_j are parameters, while ε_{ij} and ϑ_{ij} are i.i.d. normal error terms with the following distribution, $(0,0,\sigma_{\varepsilon j}^2,\sigma_{\vartheta j}^2,\rho_j\ )$. NI and I stand for "non-impaired" and "impaired" respectively.

The choice of the term "impairment" rather than disability or handicap was informed by the WHO's definition of impairment as "any loss or abnormality of psychological, physiological, or anatomical structure or function" (WHO, 1980, unpaged), a definition that most suits the impairment variable used in this study (as shown below). This is in contrast with the other terms respectively defined as "any restriction or lack (resulting from an impairment) of ability to perform an activity in the manner or within the range considered normal for a human being", and "a disadvantage for a given individual, resulting from an impairment or disability, that prevents the fulfilment of a role that is normal. . ." (WHO, 1980, unpaged).

Correction for sample selection bias adds an extra term to traditional wage-gap decompositions (Neuman & Oaxaca, 2004). For illustrative purposes, using the plausible assumption that the wage profile of the non-impaired (w_{NI}) is the non-discriminatory norm (i.e. the wage profile that would be observed in the absence of discrimination), the non-impaired versus impaired wage gap decomposition is therefore:

$$\overline{w}_{NI} - \overline{w}_I = (\bar{X}_{NI} - \bar{X}_I)\widehat{\alpha_{NI}} + \overline{X_I^{/}}(\widehat{\alpha_{NI}} - \widehat{\alpha}_I) + \left(\widehat{\theta_{NI}}\bar{\bar{\lambda}}_{NI} - \widehat{\theta}_I\bar{\bar{\lambda}}_I\right) \qquad [6.3a]$$

where the first, second and last terms on the right hand side are the respective contributions of impairment-based differences in endowment/characteristics, coefficients/prices/returns (often attributed to discrimination) and selection, to the wage-gap. Alternatively, the first and second components are also referred to as the explained and unexplained components respectively (Beblo et al., 2003). For the selection term, λ is the inverse of the Mill's ratio. Alternatively, a three-part decomposition of the non-selection part of the wage differential into characteristics, prices (or discrimination) and an interaction term is possible. It is important to acknowledge that though differences in the returns or the unexplained component is used here to denote discrimination, it also incorporates the effect of all unobserved variables to the wage gap (Jann, 2008). Additionally, even differences in observed characteristics may exist due to prior discrimination (Beblo et al., 2003). As a result, the returns component may over- or under-estimate discrimination depending on which effect dominates (this point is especially important in health-related decomposition analysis, an issue I return to subsequently). Therefore, the term "discrimination" is used throughout this thesis in a loose sense to refer to the contribution of differences in returns to the total wage gap.

How to deal with the selection term in equation (6.3a) is often a complication when conducting selectivity-corrected Blinder-Oaxaca decompositions. Most authors simply netted it out from the average raw wage gap on the left hand side (Beblo et al., 2003; Boymond, Flückiger, & Silber, 1994; Reimers, 1983). But Neuman and Oaxaca (2004) pointed out that such an adjustment only yields a selectivity-corrected wage differential. As observed by Beblo et al. (2003) and Jones et al. (2006), this quantity is simply the *potential* or *offered* wage gap and not the (more policy-relevant) average differential of *observed* wages realized by actual labour market participants. In this study, I decomposed the observed wage gap, rather than the potential gap, by obtaining a separate term for the contribution of the difference in the selection terms to the observed wage gap.

The foregoing analysis only deals with the average decomposition. A useful extension of the Blinder-Oaxaca framework to deal with the distribution of residuals was proposed by Juhn, Murphy and Pierce (1993). However, a drawback of the Juhn-Murphy-Pierce (JMP) decomposition is that it is not robust to heteroscedasticity; a shortcoming that can lead to incorrect inference if the error term is not independent and normally distributed (Melly, 2005). Melly (2005) therefore proposed a heteroscedasticity-robust model based on quantile regression, that decomposes the wage gap across the entire wage distribution in a number of steps. Firstly, the conditional wage distribution is estimated via linear quantile regression and secondly, the unconditional wage distribution is estimated by integrating the conditional distribution over the range of the covariates. This follows earlier work on quantile regression-based decomposition (Machado & Mata, 2005). Similar to the JMP decomposition, differences in the wage distribution can be decomposed into characteristics, prices and residuals (i.e. unmeasured prices/returns and characteristics/quantities).

Though conceptually feasible, the estimation of the unconditional quantiles across all quantiles is computationally burdensome. Therefore, following Antonczyk et al. (2010), I avoided estimating all possible quantile regression coefficients by estimating quantile regressions for only 49 equally-spaced quantiles beginning from the second percentile. Therefore, instead of treating the quantiles (i.e. τ) as a uniformly distributed random variable on the interval, (0,1), I treated τ as uniformly distributed on the 49 evenly-spaced percentiles on the interval [0.02, 0.98].

To demonstrate the foregoing decomposition of the unconditional distribution of log wage differences between the impaired and non-impaired groups using the notation as in the Blinder-Oaxaca decomposition above (and taking the median as a measure of central tendency for illustrative purposes in line with Melly (2005)), the wage equation for each group can be written as follows:

$$w_{ij} = X_{ij}^{/}\alpha_j(0.5) + \vartheta_{ij}, \qquad j = \{NI, I\}$$

where $\alpha_j(0.5)$ is the median regression coefficient for group j. It is now possible to isolate the effects of changes in characteristics, coefficients and residuals. The τth quantile of the counterfactual distribution of wages that would have obtained among the impaired if the characteristics distribution was that of the non-impaired can be stated as follows[11]:

$$\hat{q}(\hat{\alpha}_I, X_{NI}) = \inf\left\{q: \frac{1}{N}\sum_{i=1}^{N}\sum_{h=1}^{H}(\tau_h - \tau_{h-1})1\left[X_{i,NI}\hat{\alpha}_I(\tau_h) \le q\right] \ge \tau\right\}$$

Therefore, $\hat{q}(\hat{\alpha}_I, X_{NI}) - \hat{q}(\hat{\alpha}_I, X_I)$ is brought about by differences in characteristics. To separate the returns effect from the residual effect, it can be noted that the τth quantile of the residual distribution conditional on X can be consistently estimated by $X[\hat{\alpha}(\tau) - \hat{\alpha}(0.5)]$[12]. Defining the $H \times 1$ vector, $\hat{\alpha}_{mNI,rI}$, whose hth element is denoted, $\hat{\alpha}_{mNI,rI}(\tau_h) = [\hat{\alpha}_{NI}(0.5) + \hat{\alpha}_I(\tau_h) - \hat{\alpha}_I(0.5)]$, the counterfactual distribution that would have prevailed if the median return to characteristics had been the same as the non-impaired but the residual distribution had been that of the impaired can be estimated by, $\hat{q}(\hat{\alpha}_{mNI,rI}, X_{NI})$. Thus, it is obvious that $\hat{q}(\hat{\alpha}_{mNI,rI}, X_{NI}) - \hat{q}(\hat{\alpha}_I, X_{NI})$ is due to differences in coefficients. In the same manner, $\hat{q}(\hat{\alpha}_{NI}, X_{NI}) - \hat{q}(\hat{\alpha}_{mNI,rI}, X_{NI})$ is the result of residual differences (i.e. unmeasured prices and quantities).

Combining the foregoing, a similar decomposition to the Blinder-Oaxaca decomposition can now be carried out for the entire wage distribution as follows:

$$\hat{q}(\hat{\alpha}_{NI}, X_{NI}) - \hat{q}(\hat{\alpha}_I, X_I) = \left[\hat{q}(\hat{\alpha}_{NI}, X_{NI}) - \hat{q}(\hat{\alpha}_{mNI,rI}, X_{NI})\right] + \left[\hat{q}(\hat{\alpha}_{mNI,rI}, X_{NI}) - \hat{q}(\hat{\alpha}_I, X_{NI})\right] + \left[\hat{q}(\hat{\alpha}_I, X_{NI}) - \hat{q}(\hat{\alpha}_I, X_I)\right] \qquad [6.3b]$$

where the left hand side represents differences in the wage distribution between the non-impaired and impaired; the square brackets on the right

[11] A full description of the process of deriving the unconditional quantile distribution from conditional quantiles (which yielded the following counterfactual distribution) will be distracting while not adding to new knowledge. Basically, the process consists of obtaining the conditional CDF of wages given the characteristics, X; using the law of iterated expectations to obtain the marginal distribution of wages; and inverting this marginal distribution to obtain the unconditional quantile distribution of wages for $\tau \in (0,1)$. A full derivation can be found in Melly (2005) and Angrist and Pischke (2009).

[12] The differences in coefficients across quantiles depict the distribution of the unobservable characteristics of individuals with given covariates or the conditional distribution of the dependent variable conditional on the covariates, in a quantile regression framework (Antonczyk et al., 2010).

hand side represent differences due to residuals, (median) coefficients/returns (or discrimination) and characteristics (i.e. endowment) respectively.

6.5.2 Empirical model specification

The empirical model closely follows the above theoretical model. As in Chapter 5, identification for the selectivity-based Blinder-Oaxaca decomposition was achieved by excluding household grant receipt and household size from the wage equation.

With regard to equation (6.1), the G vector (i.e. determinants of employment) comprises years of schooling, health (the CES-D10 index of depression), a quadratic in age, number of under-17 children in the household, provincial unemployment rate, location, race, marital status, gender, household receipt of government grant, and household size.

The X vector in the outcome equation – i.e. equation (6.2)- is made up of years of schooling, health (the CES-D10 index of depression), quadratic in age, location, race, gender, number of under-17 children in the household, provincial unemployment rate, marital status, number of years spent on the job (tenure), union status, occupational category, sector of employment and number of weekly hours worked. To avoid repetition, a priori expectations are as discussed in Chapter 5 as the models are virtually identical.

As already emphasized in this thesis, discrimination is used in a very loose sense as wage gaps unexplained by observable characteristics (i.e. due to differences in returns to characteristics) may also capture the effect of non-included/unobserved productive characteristics on the wage gap. This is an important issue especially with health-related decomposition, as wage gaps between the impaired and non-impaired may simply reflect the fact that the impaired are less likely to be as productive as the non-impaired and should therefore be paid less. To mitigate this problem, a number of the above-mentioned controls deserve particular mention. Firstly, education and labour market experience (the latter, proxied by age) are especially important controls that are likely to help obtain the pure discrimination effect. Moreover, it is important to realize that disability/impairment is likely to be associated with emotional distress. This is likely to be associated with reduced productivity even when the above host of characteristics have been controlled for. Therefore, controlling for the respondent's state of emotional health is likely to enable one to obtain a cleaner measure of pure discrimination. Additionally, much of wage differences between both groups is likely to arise from the fact that the impaired are likely to put in less hours at work (due to, say debilitation or time used in seeking health care). To the extent that wage differences arise nontrivially from differences in time spent working, controlling for hours worked (as done in this study) mitigates the non-discrimination component of the coefficient/returns component.

6.5.3 Data and descriptive statistics

In this section, I describe the variables used in the regression analysis, set up the empirical model and provide a description of the data used in this study. Unless otherwise stated, the sample in both waves (i.e. wave 1 and wave 3) consists of African and coloured adults aged 20-56 years in wave 1. This is to ensure the comparability of results for both waves. The racial restriction above was necessitated by the small number of impaired whites and Asians/Indians (according to the definition of impairment). Also, only employees in wage employment were considered in the wage equation (i.e. self-employed and casual workers were excluded).

Impairment was defined as a dummy variable which equals one if the respondent suffered from any of the following conditions: memory loss, tuberculosis, stroke, physical handicap and psychological/psychiatric disorder. The impairment variable can be considered similar to Johnson and Lambrinos' (1985) measure of handicap and in like manner, I defined impairment to include those conditions likely to evoke prejudice in the labour market. However, my measure contains fewer health conditions than in Johnson and Lambrinos given the need to not include too many conditions to the extent that one loses sense of what is being evaluated (the handicap variable in Johnson and Lambrinos included seven conditions, "total deafness, inability to read ordinary size print with glasses, blindness, partial or complete paralysis, convulsive disorders, distortion of limbs or spine, and mental illness" (Johnson & Lambrinos, 1985, p. 266)).

A potential problem with this measure of impairment is the identical treatment of the effect of the union and intersection of these health conditions, i.e. it assumes that the effect of suffering from one condition is identical to that of suffering from another condition or even having both impairment conditions concurrently. This is a common problem with dichotomous variables made up of multiple measures. It would have been ideal to ascertain the amount of wage discrimination (if any) associated with only one condition at a time. However, the choice of this "multiple conditions" impairment variable was necessitated by the small sample of individuals suffering from each condition. For instance, the number of respondents in the estimation sample suffering from memory loss, tuberculosis, stroke, physical handicap and psychological/psychiatric disorder in wave 1 (wave 3) were 109(142), 65(29), 9(6), 18(4), and 11(3) respectively. Thus, it is virtually impractical to obtain robust estimates from models based on each of these individual conditions. However as subsequently shown, the results were robust to the inclusion of only memory loss and tuberculosis (the two conditions with the largest sub-samples of "impaired" wage earners).

The descriptive statistics is presented in Table 6.1 below.

Table 6.1: Descriptive statistics

Variable	N	Mean	Std.Dev.[a]	N	Mean	Std.Dev.
		Wave 1			Wave 3	
employee[†]	7122	0.47	0.5	5437	0.58	0.5
impaired	9496	0.07	0.3	7830	0.08	0.3
real wage	3156	3993.00	6053.0	3054	4593.03	5353.0
schooling	11321	9.04	3.7	8778	9.21	3.8
cesd10	9479	8.34	4.8	7824	7.13	4.4
age	11400	34.29	10.0	8795	38.95	10.2
num. children[††]	11400	1.78	2.0	8795	1.45	1.7
prov. unemp[‡]	11400	23.58	3.0	8795	25.55	3.5
uf only[‡‡]	11400	0.50	0.5	8775	0.50	0.5
African	11400	0.89	0.3	8795	0.90	0.3
coloured	11400	0.11	0.3	8795	0.10	0.3
married	9481	0.43	0.5	7828	0.43	0.5
male	11400	0.48	0.5	8775	0.47	0.5
grant	11400	0.55	0.5	8775	0.57	0.5
household size	11400	4.71	3.3	8775	4.76	3.2
tenure	3088	6.87	7.7	3030	7.70	8.7
union	3076	0.34	0.5	3047	0.35	0.5
manprof[¥]	3115	0.19	0.4	3014	0.21	0.4
tertiary	2916	0.50	0.5	2919	0.63	0.5
hours	2880	40.58	14.8	3001	41.04	13.5

Source: NIDS wave1 and wave 3; author's calculations. Sample is restricted to 20-56 year-old respondents in wave 1. Both samples corrected for survey design, national representativeness and non-random attrition; [a] standard deviation; [†]dummy variable (=1 if respondent is an employee; 0 otherwise); [††]number of under-17 children in the household; [‡]provincial unemployment rate; [‡‡]dummy variable (=1 if respondent resides in urban formal area; 0 otherwise); [¥]dummy variable (=1 if respondent belongs to the managerial/professional occupational category; 0 otherwise)

As shown in Table 6.1, the various means and proportions did not change much for most variables over the four-year period. The proportion adjudged impaired remained virtually unchanged across waves while the proportion in wage employment increased from 47% to 58%. Average individual monthly real wage increased by 15%. As expected in a highly unequal country like South Africa, the respective wage standard deviations were high though declining: R6053 and R5353. Average years of completed schooling in wave 1 (wave 3) was approximately 9 in both waves while the rate of unionization remained virtually unchanged. Furthermore, 19% (21%) of workers belonged to the managerial/professional occupational category in wave 1 (wave 3).

Finally, average number of years so far spent on current job (i.e. tenure) were 6.9 and 7.7 in waves 1 and 3 respectively.

Table 6.2 depicts impairment-related gaps across key variables in both waves. Real wage, years of schooling, employment probability as well as probabilities of being a manager/professional and a union member were higher among the non-impaired relative to the impaired across waves. Also, the real wage gap widened across waves. Moreover, the impaired were significantly older than the non-impaired across waves. These results conformed to expectations.

On the whole, there is at least descriptive evidence suggesting impairment-related disparities in individual characteristics in general and wages in particular. However, it cannot be suggested that such disparities are due to discrimination till more rigorous analysis is conducted. Ascertaining the discrimination component of the foregoing impairment-related wage gaps as well as its pattern over time is the focus of the next section.

Table 6.2: Relationship between impairment status and key variables among working age Africans and coloureds: two-sample tests of mean and proportional differences (p value of test statistic in parenthesis)

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
			Wave 1				Wave 3	
Variable	Sample size (Non-impaired/impaired)	Non-impaired	Impaired	Mean difference (3-4)	Sample size (Non-impaired/impaired)	Non-impaired	Impaired	Mean difference (7-8)
real wage	2979/173	3248.92	2043.61	1205.31*** (0.0)	2896/158	4033.95	2719.74	1314.22*** (0.0)
age	8713/783	35.0	40.4	-5.5*** (0.0)	7109/721	39.4	44.6	-5.3*** (0.0)
schooling	8697/782	8.5	6.1	2.4*** (0.0)	7102/718	8.8	6.7	2.0*** (0.0)
hours	2721/156	40.9	39.6	1.3 (0.13)	2846/155	41.2	39.3	2.1** (0.03)
employee†	6640/476	0.45	0.36	0.06*** (0.0)	5076/359	0.57	0.44	0.13*** (0.0)
manprof††	2940/172	0.18	0.11	0.07** (0.01)	2857/157	0.20	0.15	0.05* (0.07)
union	2905/169	0.30	0.24	0.07** (0.03)	2889/158	0.32	0.26	0.06** (0.048)

Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1; statistics corrected for complex survey design and national representativeness; † dummy variable (=1 if respondent is an employee; 0 otherwise); ††dummy variable (=1 if respondent belongs to managerial/professional occupational category; 0 otherwise)

6.5.4 Econometric results

Blinder-Oaxaca decomposition

Table 6.3 displays impairment-related wage decompositions for both OLS and selection-corrected wage regressions in waves 1 and 3 respectively using the (more realistic) wage profile of the non-impaired as the non-discriminatory norm. These estimates were similar to estimates obtained from using the wage profile from a pooled regression of both impairment groups (as suggested by Neumark (1988)), as well as that of the impaired group as the non-discriminatory norm (both are available on request).

Table 6.3: Impairment-related Blinder-Oaxaca log wage decompositions

1	2	3	4	5	6	7	8	
Non-discriminatory norm	N/N1/N2	Raw gap	Xteristics†	Returns	Selection	Estimated gap	Percentage contribution to estimated gap	
							Xteristics	Returns
			Wave 1					
OLS estimates	10686/2379/139	0.51	0.31***	0.15**		0.46***	67.4	32.6
			(0.08)	(0.07)		(0.10)		
Selection-corrected estimates	11258/2379/139	0.51	0.24***	0.79***	-0.75	1.03***	23.3	76.7
			(0.07)	(0.20)	(0.55)	(0.21)		
			Wave 3					
OLS estimates	7983/2661/145	0.42	0.24***	0.22***		0.47***	51.1	46.8
			(0.06)	(0.07)		(0.08)		
Selection-corrected estimates	8734/2661/145	0.42	0.19***	0.69***	-0.51	0.88***	21.6	78.4
			(0.05)	(0.25)	(0.46)	(0.26)		

Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1; statistics corrected for complex survey design and national representativeness; standard errors of difference between selection terms obtained via bootstrapping with 400 replications; non-impaired wage profile regression serves as non-discriminatory norm; N ⇒sample size; N1 ⇒number of observations in non-impaired group; N2 ⇒ number of observations in impaired group; †characteristics/endowment

From Table 6.3, the estimated wage gap was higher in selectivity-corrected decompositions relative to OLS-based decompositions in both waves. While the estimated gap obtained from the OLS regression was 0.46 (0.47) in wave 1 (wave 3), it was 1.03 (0.88) respectively for the selection-corrected decomposition. In terms of the portion of the estimated gap contributed by both the characteristics and returns components, column 8 shows that the characteristics (returns/discrimination) contributions were 67% (33%) and 23% (77%) respectively for the OLS and selection-corrected decompositions in wave 1. In wave 3, characteristics (returns) contributions were 51% (47%) for OLS and 22% (78%) for the decomposition based on the sample-selection model. Thus, the percentage contribution of the discrimination component increased across waves in both OLS and selection-corrected models (though the latter remained virtually stable with only a one percentage point increase). Moreover, the magnitude of the wage gap contributed by the endowment (i.e. characteristics) component decreased over the waves. However, these temporal changes in both the endowment and discrimination components as well as in the aggregate estimated gaps were not statistically significant ($p>0.1$). Furthermore, the negative selection effect recorded in Table 6.3 is consistent with a lot of Blinder-Oaxaca decompositions (albeit mostly gender-based), implying that the offered wage gap exceeded the observed wage gap (see e.g. Ogloblin, 1999). It was however, not statistically significant in both waves.

With regard to previous evidence, Baldwin and Johnson (1994) observed that discrimination explained 47% of the total wage gap against the handicapped in the US while the figure was 48.2% against disabled men in the US (Kidd et al., 2000). Similarly, for the exogenous treatment of health as in this study, Madden (2004) maintained that discrimination explained 30% of the wage gap against men with long term illness. Using the wage profile of the non-disabled as the non-discriminatory norm in a selectivity-corrected framework (similar to this study), Jones et al, (2006) observed that discrimination accounted for 65.9% (51.4%) of the wage gap between work-limiting disabled workers relative to the non-disabled in 1997 (2003) among

British men. For women, the figures were 56.2% and 75.7% in 1997 and 2003 respectively. However, the percentage contribution of discrimination to the wage gap was higher between the non work-limiting disabled and the non-disabled, though the absolute wage gaps were smaller than between the two previous groups. However, a drawback of the Jones et al. (2006) study is that the statistical significance of these wage gaps were not reported. Unfortunately, I am unaware of developing country estimates of impairment-related wage gaps obtained using similar methodology as used in this chapter, but as these estimates show, the component of impairment/disability-related wage gaps accounted for by discrimination in South Africa is similar in proportional terms to much of what obtained in developed countries.

To ascertain whether the estimates were robust to the components of the impairment variable, I re-estimated the above models with an impairment variable made up of memory loss and tuberculosis, the two sub-components with the largest number of the impaired. The results (Table A6.1 in the Appendix) were robust to this restriction and a comparison of Table 6.3 and Table A6.1 indicate that most of the results were mainly driven by these two health conditions. As I mentioned earlier, it would have been desirable to ascertain whether discrimination was present for each impairment variable but data constraints prevented such analysis.

Given the possibility of different dynamics relating to wage discrimination in public and private sectors, it would have been desirable to conduct the analysis separately for both sectors. Unfortunately, the data does not permit such an analysis.

Table 6.4 below depicts the contributions of individual regressors to the wage gap. Most of the regressors did not significantly contribute to both the explained and unexplained gaps individually across waves. However, years of schooling and occupational class significantly contributed to the explained gap in both waves. Union membership only contributed in wave 3. From the results, it is clear that education was the single most important determinant of explained impairment-related wage gaps as measured in this study, followed by occupational category.

Table 6.4: Contributions of variable groups to wage gap

	(1)	(2)	(3)	(4)
	Wave 1		Wave 3	
VARIABLES	explained	unexplained	explained	unexplained
schooling	0.10***	-0.15	0.10***	0.39*
	(0.04)	(0.17)	(0.03)	(0.23)
age	-0.00	-0.17	-0.00	0.42
	(0.01)	(1.20)	(0.01)	(1.29)
uf only†	0.03	-0.04	-0.00	-0.20
	(0.02)	(0.08)	(0.00)	(0.13)
coloured	0.00	-0.02	-0.01	0.02
	(0.00)	(0.02)	(0.01)	(0.02)
male	0.01	0.01	0.02	0.05
	(0.02)	(0.07)	(0.02)	(0.07)
num. children††	-0.00	0.12**	-0.00	0.12*
	(0.01)	(0.05)	(0.00)	(0.06)
prov. unemp‡	-0.00	-0.57	-0.00	0.97
	(0.01)	(0.66)	(0.00)	(0.65)
married	0.01	-0.12**	0.00	-0.08
	(0.01)	(0.05)	(0.00)	(0.07)
tenure	-0.00	-0.12*	-0.01	-0.08
	(0.01)	(0.07)	(0.02)	(0.08)
union	0.01	-0.02	0.04**	0.04
	(0.02)	(0.05)	(0.02)	(0.04)
manprof‡‡	0.06***	0.00	0.04**	-0.00
	(0.02)	(0.02)	(0.02)	(0.03)
tertiary	0.01	-0.02	0.00	-0.04
	(0.01)	(0.06)	(0.01)	(0.10)
hours	0.00	0.18	0.00	-0.20
	(0.00)	(0.18)	(0.00)	(0.21)
cesd10	0.02	-0.09	0.01	0.18
	(0.01)	(0.13)	(0.01)	(0.17)
constant		1.78		-0.90
		(1.29)		(1.54)
Observations	11,258	11,258	8,734	8,734

Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1; †dummy variable (=1 if respondent resides in urban formal area; 0 otherwise); ††number of under-17 children in the household; ‡provincial unemployment rate; ‡‡dummy variable (=1 if respondent belongs to the managerial/professional occupational category; 0 otherwise)

A general caveat in explaining detailed decomposition results pertaining to categorical covariates (with more than two categories) in the Blinder-Oaxaca model is that the choice of the omitted category may influence the interpretation of the results if care is not taken. Additionally, it becomes difficult to differentiate between the portion of the unexplained part attributed to group membership (captured by intercept differences) and the portion attributed to differences in the coefficient of the base category (Fortin, Lemieux, & Firpo, 2011). However, this is not an issue in this thesis as all categorical covariates are binary.

The underlying regression results that yielded the above decomposition results are shown in the Appendix in Table A6.2 (for OLS) and Table A6.3 and Table A6.4 (for selection-corrected results in wave 1 and wave 3 respectively). The results largely conformed to a priori expectations for both the impaired and the non-impaired. For the participation equation in the sample selection model, education and being male were positively associated with being employed, while household grant receipt and a higher depression score had a negative association with employment. Age had a quadratic relationship with employment. Also, job tenure, education, unionization, being in the managerial/professional occupational class, being male and living in urban formal areas were positively associated with higher wages. As the selection-corrected results indicate, there existed negative sample selection bias among the non-impaired while such evidence was lacking among the impaired. The plausible explanation of such a negative selection effect, which was provided in Chapter 5, also applies here. Furthermore, the non-significance of the selection variable among the impaired may be due to their small sample size.

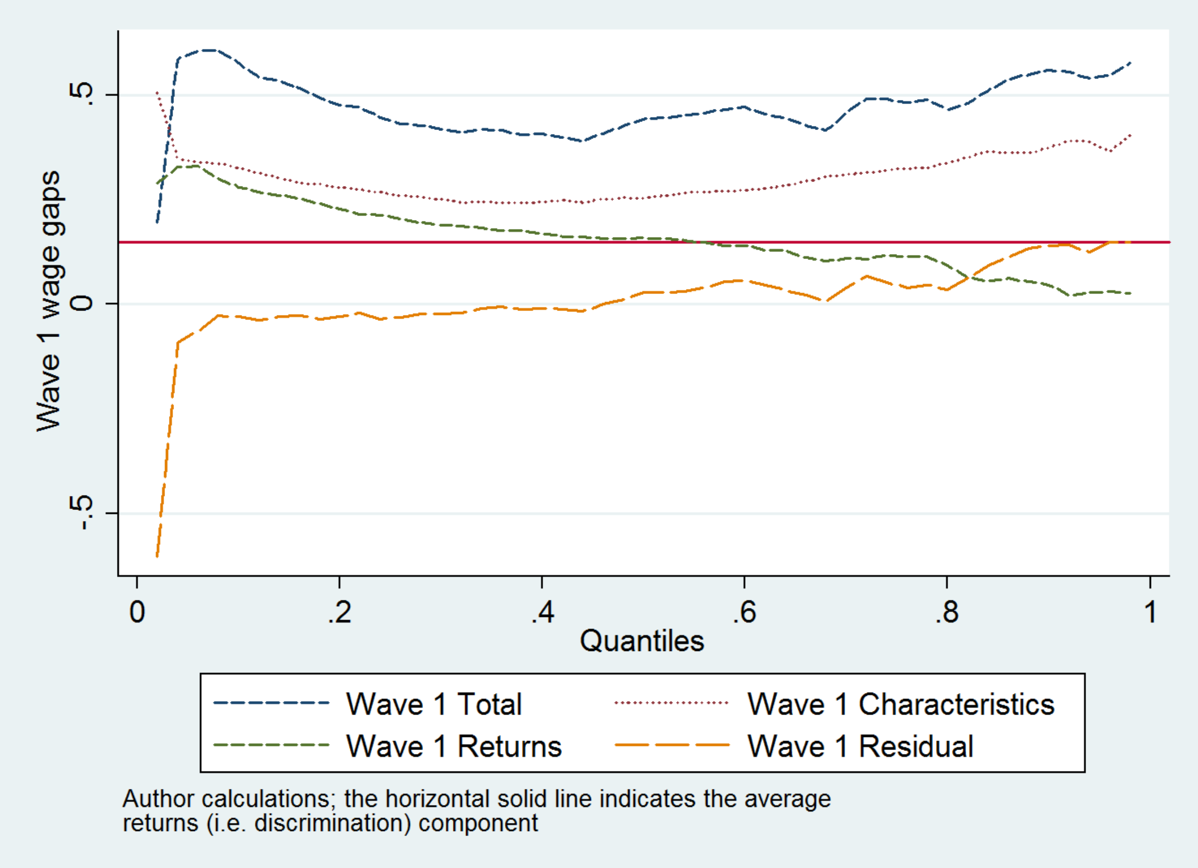
With regard to magnitudes, for the selection-corrected wage regressions, an additional year spent on the job was associated with 2% wage increase while one additional year of schooling was associated with 7-8% wage increase. Union premium ranged between 32% and 43%. Substantial union premium has also been previously found among Africans where union membership among African males in the bottom wage decile was associated with wages that were 145% higher than those of their non-unionized counterparts (Schultz & Mwabu, 1998b). Apparently, such a premium has persisted over time. Managers/professionals earned about 42-51% higher than those in semi-skilled and elementary occupations. A gender-based analysis was not conducted due to the small sample size of the impaired. The small sample size for the impaired also resulted in a smaller number of statistically significant variables in the impaired regressions relative to the non-impaired.

Decomposition of differences in wage distributions

As earlier noted, a drawback of the Blinder-Oaxaca technique is that it is not very informative regarding other parts of the (wage) distribution other than the mean. The main purpose of this section is to analyse the estimated distribution of the wage gap as well as its components across the wage distribution in each wave. Following equation (6.3b), I illustrate the differential impairment-related decomposition results for workers at 49 quantiles of the log wage distribution in Figure 6.1 and Figure 6.2 (wave 1 and wave 3 respectively) for each of the components of the log wage gap[13]. The numerical contributions of the different components for only select quantiles (precisely nine deciles) are presented in Table A6.5 in the appendix. Comparing Table A6.5 and Table 6.3, the total estimated gap for the median was virtually identical to the mean total estimated gap in the Blinder-Oaxaca model. Following Melly (2005), the standard errors were computed via a bootstrap procedure with 100 replications.

[13] Selectivity bias was not controlled for in these figures.

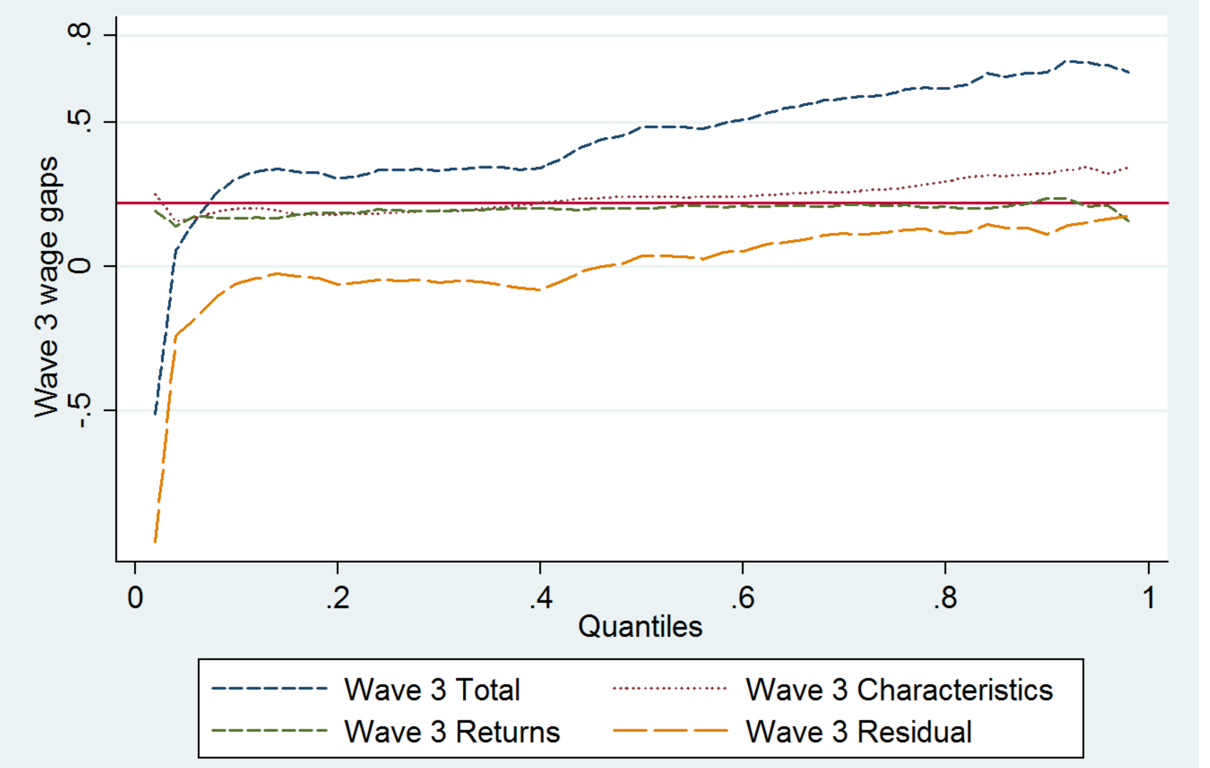
Figure 6.1: Quantile regression-based decomposition of impairment-related differences in distribution (wave 1: 2008)



As Figure 6.1 shows, the degree of variation in the total gap as well as its various components differed slightly. The total estimated gap as well as the portion contributed by differences in characteristics was slightly u-shaped for most part of the distribution, indicating similar magnitude of the wage gap at both ends of the wage distribution which was slightly higher than what obtained around the median. The total gap was 0.58 and 0.56 among respondents in the 10th and 90th percentiles respectively (i.e. above the average of 0.46 from the Blinder-Oaxaca decomposition results) but a low of 0.41 at the 40th percentile. Differences due to returns gently decreased across the wage distribution as one moved up the wage distribution. Therefore in terms of magnitude, Figure 6.1 indicates that discrimination was more acute among low earners relative to other parts of the wage distribution. This is clearly shown by comparing the returns curve and the solid horizontal average returns line in Figure 6.1 (wage gap=0.15, obtained from the above Blinder-Oaxaca decomposition). However, the null hypothesis of constant quantile effects across the wage distribution for the returns/discrimination component could not be rejected based on the Kolmogorov-Smirnov test statistic (p>0.10). This suggests homogeneity of the absolute contribution of the returns/discrimination component across the wage distribution. Finally, the distribution of differences in the residual contributed least in explaining the wage gaps for most portions of the distribution.

In wave 3 (see Figure 6.2 below), differences in returns accounted for a substantial part of the estimated wage gap across the wage distribution. While its relative contribution declined among higher earners as in wave 1, its absolute contribution generally remained flat across the wage distribution. This suggests homogeneity of the discrimination component, a claim supported by failure to reject the null hypothesis of constant quantile effects of the returns component as in wave 1 (p>0.10).

Figure 6.2: Quantile regression-based decomposition of impairment-related differences in distribution (wave 3: 2012)



Author calculations; the horizontal solid line indicates the average returns (i.e. discrimination) component

As shown in the Blinder-Oaxaca results in Table 6.3, the proportion of the total estimated wage gap contributed by the returns/discrimination component increased between wave 1 and wave 3 using both OLS and sample selection specifications. The above distributional decomposition supports this finding as the proportion of total estimated wage gaps explained by the returns/discrimination component was higher in wave 3 relative to wave 1 in all but one of the estimated 49 quantiles. Even in terms of numerical magnitude, the wave 3 returns component generally exceeded their wave 1 counterparts in 35 quantiles (6 deciles in Table A6.5). This implies that impairment-related wage discrimination not only increased in relative terms between 2008 and 2012 on the average (based on the Blinder-Oaxaca decomposition), most parts of the wage distribution also experienced a temporal increase in the returns component in both absolute and relative terms among African and coloured South African employees (based on the distributional decomposition).

Pinpointing the exact reason(s) for such proportional increase in impairment-related wage discrimination between 2008 and 2012 is not easy as many factors might have been at play. However, it may not be unconnected with poor growth in post-2008 recession South Africa (for instance, data from the South African Reserve Bank shows that GDP per capita growth rate increased from 0.2% in 1999 to 4.1% in 2007 but declined to 1.3% in 2012 (South African Reserve Bank, undated). This may have placed a lot of stress on the labour market to the extent that impairment was viewed in more negative light relative to the pre-2008 period when the economy experienced a boom.

6.6 CONCLUSION

This chapter set out to ascertain the magnitude of impairment-related wage discrimination (i.e. unexplained wage gaps), the temporal nature of such gaps as well as the heterogeneity of discrimination over the wage distribution in the South African labour market. From the results, non-trivial impairment-related wage discrimination existed among African and coloured employees in both 2008 and 2012. Average returns to characteristics (loosely termed discrimination) increased over time as a percentage of total estimated wage gaps when using both ordinary least squares and sample selection models as the basis for the Blinder-Oaxaca decompositions. This was similar to more nuanced findings from distributional decompositions, as the ratio of the contribution of returns to the total gap was higher in 2012 relative to 2008 in most deciles of the wage distribution. Moreover, the discrimination component of total estimated wage gaps was largely homogenous across the wage distribution in each wave. Finally, education was the most important factor in determining explained wage gaps while occupational class and union membership also played important roles (the latter only in wave 3).

CHAPTER 7

CONCLUSION

This thesis has examined a key issue in the economic growth literature: the role of health in the determination of labour market outcomes. Specific issues investigated were the impact of health on labour force participation; the gradient between physical, psychological and general health on the one hand, and wages on the other; and the existence of impairment-related differences in returns to human characteristics (i.e. wage discrimination). These issues were examined both descriptively and more formally using various econometric methods like instrumental variables, quantile regression, Blinder-Oaxaca decomposition and quantile regression-based distributional decomposition.

Several insights were uncovered in the analysis. Firstly, it was found that better health (measured as respondents' self-reports of their overall health status) had a positive impact on labour force participation at the extensive margin irrespective of whether the broad or strict measure of LFP was used. In numerical terms, average treatment effects ranged from 20 to 23%, treatment effect on the treated, 29 to 33%, and local average treatment effect, 23 to 26%. Even if one is not convinced about the exogeneity of the instruments used for identification, there was still evidence of a positive association between both health and LFP in both cross-sectional and temporal settings.

Furthermore, positive and nontrivial gradients were uncovered between better health (physical, psychological and general) and wages for African and coloured employees. This was largely true for the average worker as well as for those at the bottom quartile, median and top quartile of the earnings profile. The gradients for physical and psychological health (as well as the median of the wage distribution for general health) were persistent even four years after report of the health condition. Also, the physical health gradient was steeper for males relative to females, possibly due to greater demand for physical strength for jobs mostly done by men.

Again, it was found that substantial impairment-related differences in returns to characteristics (loosely termed wage discrimination) existed among Africans and coloureds in South Africa whether or not sample selection was accounted for. It also increased as a proportion of the total estimated wage gap between 2008 and 2012. More nuanced analysis showed that this temporal proportional increase occurred for all but one of the estimated quantiles of the wage distribution. Also in terms of magnitude, the discrimination component in wave 3 exceeded its wave 1 counterpart in most of the estimated quantiles.

For the various controls, I found that males were more likely than females to be labour force participants. Also, household grant receipt was associated with reduced LFP probability (possibly due to higher reservation wages of grant-receiving household members) while education and age were associated with increased LFP. Location also played an important role in determining LFP as living in traditional authority areas was associated with reduced LFP relative to residing in other areas. Furthermore, the presence of at least one employed male in the household was associated with increased female LFP probability, while marriage/cohabitation was negatively (positively) associated with female (male) participation. For wage determination, education, occupational class, gender and industry were very important factors. Finally, education and occupational class mostly contributed to the explained part of impairment-related wage gaps among Africans and coloureds between 2008 and 2012.

LIMITATIONS AND FURTHER RESEARCH

Though this thesis has uncovered important issues regarding the relationship between health and the labour market, it is by no means perfect. It was limited by time and data constraints. Therefore, a number of shortcomings that should form the basis for future research are apparent.

Firstly, an important phenomenon in the labour supply discourse is the issue of the impact of health on presenteeism, i.e. on-the-job productivity loss (Goetzel et al., 2004). Goetzel et al. have demonstrated that different health conditions like heart disease, depression and hypertension are associated with substantial absenteeism and presenteeism costs in the United States. This study could not investigate the health-presenteeism relationship due to data constraints.

Additionally, examining the health-wage gradient for the informal sector and the self-employed will throw more light on the relationship for the entire labour market. In this thesis, I focused on the gradient among wage employees only given the small number of individuals in the sample who are either self or informally employed. This is due to the small size of the informal sector in South Africa relative to other African countries (Kingdon & Knight, 2007). The above is also true for impairment-related wage discrimination among those in informal/self-employment.

Furthermore, it would be interesting to link the discrimination results to theory, i.e. to test whether, say, prejudice or information asymmetry is the likely underlying or dominant cause of impairment-related wage discrimination. For instance, a possible means of testing prejudice-based discrimination exploits the notion that discriminatory wage differentials against the impaired should vary according to the intensity of prejudice against specific impairments if indeed prejudice is a key determinant of impairment-related wage discrimination (Baldwin & Johnson, 2006). Such an exercise will involve estimating the impairment-related wage discrimination associated with different ailments that are often associated with varying degrees of prejudice and functional limitation. For instance, epilepsy is often associated with more intense prejudice but less work limitation relative to arthritis and vice versa. In this case, evidence of more impairment-related wage discrimination among epilepsy sufferers relative to arthritis patients may be suggestive of prejudice-based discrimination. Unfortunately, such analysis was not practical in this thesis given the small sample of the impaired for separate impairment types as earlier mentioned. A larger survey in the future may enable one conduct such analysis.

Finally, most of the econometric analyses in this thesis were confined to Africans and coloureds. This therefore limited the generalizability of the findings especially given that the various racial groups are different in many respects especially in their labour market experiences due to apartheid. In the future, analyses that encompass all racial groups may be possible with larger surveys with the richness of NIDS.

APPENDIX

Table A4.1: Non-health ("traditional") determinants of LFP (marginal effects)

	(1)	(2)	(3)	(4)	(5)	(6)
		Probit			LPM	
Variables	Total	Female	Male	Total	Female	Male
grant	-0.07***	-0.06***	-0.08***	-0.07***	-0.06***	-0.08***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)
matric	0.12***	0.16***	0.06**	0.11***	0.16***	0.05**
	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.03)
age26-30	0.14***	0.18***	0.09***	0.14***	0.19***	0.10***
	(0.02)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)
age31-35	0.14***	0.16***	0.11***	0.14***	0.17***	0.11***
	(0.02)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)
age36-40	0.18***	0.20***	0.14***	0.18***	0.21***	0.13***
	(0.02)	(0.03)	(0.04)	(0.02)	(0.03)	(0.03)
age41-45	0.13***	0.13***	0.11**	0.13***	0.14***	0.11***
	(0.02)	(0.03)	(0.05)	(0.02)	(0.03)	(0.04)
age46-50	0.08***	0.11***	0.02	0.08***	0.11***	0.04
	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)
age51-60	-0.04	-0.03	-0.07**	-0.05*	-0.04	-0.06*
	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)
rural formal	0.06**	0.01	0.11***	0.07***	0.02	0.12***
	(0.03)	(0.03)	(0.04)	(0.03)	(0.04)	(0.03)
urban formal	0.11***	0.11***	0.09***	0.11***	0.12***	0.09***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)
urban informal	0.06**	0.09***	0.02	0.07**	0.10***	0.03
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)
African	0.02	0.04	-0.02	0.02	0.04	0.00
	(0.04)	(0.05)	(0.05)	(0.04)	(0.05)	(0.04)
coloured	-0.01	0.06	-0.09	0.00	0.06	-0.06
	(0.05)	(0.06)	(0.06)	(0.04)	(0.05)	(0.05)
prov. unemp†	0.00	-0.00	0.00	0.00	-0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
married††	0.03*	-0.03	0.10***	0.03*	-0.03	0.09***
	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)
male	0.12***			0.11***		
	(0.01)			(0.01)		
num. children‡	-0.00	-0.00	-0.00	-0.01	-0.01	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
household size	-0.01*	-0.01	-0.01	-0.01	-0.01	-0.01
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)
employed male‡‡		0.04**			0.04**	
		(0.02)			(0.02)	
constant				0.49***	0.49***	0.60***
				(0.07)	(0.09)	(0.09)
F-stat	38.0	20.4	12.8	52.7	28.6	18.3
R^2 adjusted				0.12	0.11	0.10
N	9791	5832	3959	9791	5832	3959

Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1; estimates corrected for complex survey design and non-random attrition; †provincial unemployment rate; ††married/cohabiting; ‡number of under-17 children in household; ‡‡household has at least one employed male

Table A4.2: First stage estimates: IV-LPM

Variables	(1) Total	(2) Female	(3) Male
joint pain	-0.25***	-0.25***	-0.23***
	(0.03)	(0.03)	(0.06)
memory loss	-0.17***	-0.15***	-0.20***
	(0.04)	(0.04)	(0.06)
grant	-0.02***	-0.02*	-0.02
	(0.01)	(0.01)	(0.01)
matric	0.03***	0.04***	0.02*
	(0.01)	(0.01)	(0.01)
age26-30	-0.02*	-0.02	-0.01
	(0.01)	(0.02)	(0.01)
age31-35	-0.06***	-0.07***	-0.05***
	(0.02)	(0.03)	(0.02)
age36-40	-0.05***	-0.06***	-0.03**
	(0.01)	(0.02)	(0.01)
age41-45	-0.08***	-0.07***	-0.11***
	(0.02)	(0.02)	(0.03)
age46-50	-0.11***	-0.11***	-0.12***
	(0.02)	(0.03)	(0.02)
age51-60	-0.14***	-0.16***	-0.11***
	(0.02)	(0.03)	(0.02)
rural formal	-0.06**	-0.08**	-0.02
	(0.02)	(0.03)	(0.03)
urban formal	-0.01	-0.03*	0.01
	(0.01)	(0.02)	(0.01)
urban informal	-0.02*	-0.04*	-0.00
	(0.01)	(0.02)	(0.02)
African	-0.02	-0.03	-0.01
	(0.02)	(0.03)	(0.03)
coloured	-0.03	-0.03	-0.03
	(0.02)	(0.03)	(0.03)
prov. unemp†	-0.00**	-0.00	-0.01**
	(0.00)	(0.00)	(0.00)
married††	0.04***	0.05***	0.02
	(0.01)	(0.01)	(0.01)
male	0.02***		
	(0.01)		
num. children‡	0.00	-0.00	0.00
	(0.00)	(0.01)	(0.01)
household size	0.00	0.00	-0.00
	(0.00)	(0.00)	(0.00)
employed male‡‡		-0.01	
		(0.01)	
constant	1.08***	1.05***	1.13***
	(0.04)	(0.05)	(0.06)
R^2 adjusted	0.13	0.14	0.10
N	9795	5836	3956
F-stat	27.2	20.8	7.6

Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1; estimates corrected for complex survey design and non-random attrition; †provincial unemployment rate; ††married/cohabiting; ‡number of under-17 children in household; ‡‡household has at least one employed male

Table A5.1: OLS results of health-wage relationships

	(1)	(2)	(3)	(4)	(5)	(6)
		Wave 1			Wave 3	
Variables						
lnbmi	0.36***			0.33***		
	(0.07)			(0.06)		
lncesd10		-0.09***			-0.06***	
		(0.02)			(0.02)	
sah			0.22***			0.29***
			(0.04)			(0.07)
tenure	0.01***	0.01***	0.01***	0.01***	0.01***	0.01***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
matric	0.62***	0.63***	0.63***	0.59***	0.60***	0.59***
	(0.04)	(0.04)	(0.04)	(0.03)	(0.03)	(0.03)
age	0.02*	0.03**	0.03**	0.02	0.02*	0.02*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
age^2	-0.00	-0.00*	-0.00**	-0.00	-0.00*	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
union	0.51***	0.52***	0.53***	0.47***	0.48***	0.48***
	(0.04)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)
manprof†	0.39***	0.42***	0.41***	0.49***	0.50***	0.51***
	(0.05)	(0.05)	(0.05)	(0.04)	(0.04)	(0.04)
tertiary	0.19***	0.19***	0.18***	0.20***	0.21***	0.21***
	(0.04)	(0.04)	(0.04)	(0.03)	(0.04)	(0.04)
formal loc††	0.17***	0.18***	0.18***	0.14***	0.15***	0.15***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
coloured	0.01	-0.05	-0.04	0.01	-0.02	0.00
	(0.05)	(0.05)	(0.05)	(0.04)	(0.05)	(0.04)
hours	0.01***	0.00***	0.00***	0.01***	0.01***	0.01***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
male	0.34***	0.28***	0.28***	0.38***	0.32***	0.32***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
prov. unemp‡	-0.01	-0.01	-0.01	-0.00	-0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)
married	0.07**	0.07**	0.07**	0.08***	0.09***	0.09***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
num. children‡‡	-0.03***	-0.02**	-0.02**	-0.03***	-0.02**	-0.02***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
constant	5.01***	6.26***	5.89***	5.29***	6.39***	5.94***
	(0.36)	(0.26)	(0.26)	(0.34)	(0.29)	(0.28)
N	2,189	2,518	2,505	2,747	2,806	2,805
R^2 adjusted	0.48	0.49	0.49	0.48	0.48	0.49
F-stat	127.9	133.2	132.1	125.9	130.9	134.5

Cluster-robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; † dummy variable (=1 if respondent is in the managerial/professional occupational category); †† dummy variable (=1 if respondent resides in a formal location; 0 otherwise); ‡ provincial unemployment rate; ‡‡ number of under-17 children in the household

Table A5.2: Heckman selection results of health-wage gradient in wave 1

	(1)	(2)	(3)	(4)	(5)	(6)
Variables	lnrealwage	employee	lnrealwage	employee	lnrealwage	employee
lnbmi	0.26***	0.28***				
	(0.07)	(0.08)				
lncesd10			-0.05**	-0.10***		
			(0.02)	(0.03)		
sah					0.11**	0.23***
					(0.05)	(0.05)
tenure	0.01***		0.01***		0.01***	
	(0.00)		(0.00)		(0.00)	
matric	0.47***	0.35***	0.47***	0.37***	0.47***	0.36***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
age	-0.03**	0.12***	-0.02*	0.12***	-0.02*	0.12***
	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
age^2	0.00*	-0.00***	0.00*	-0.00***	0.00*	-0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
union	0.49***		0.51***		0.51***	
	(0.04)		(0.03)		(0.03)	
manprof†	0.39***		0.42***		0.41***	
	(0.05)		(0.05)		(0.05)	
tertiary	0.18***		0.18***		0.18***	
	(0.04)		(0.04)		(0.04)	
formal loc††	-0.04	0.50***	-0.04	0.50***	-0.04	0.50***
	(0.05)	(0.06)	(0.05)	(0.06)	(0.05)	(0.06)
coloured	-0.05	0.21**	-0.11*	0.18**	-0.10	0.19**
	(0.06)	(0.08)	(0.06)	(0.08)	(0.06)	(0.08)
hours	0.00***		0.00***		0.00***	
	(0.00)		(0.00)		(0.00)	
male	0.18***	0.39***	0.15***	0.32***	0.16***	0.31***
	(0.04)	(0.04)	(0.03)	(0.04)	(0.03)	(0.04)
prov. unemp‡	0.00	-0.03***	0.00	-0.02**	0.00	-0.03***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
married	0.04	0.03	0.04	0.05	0.04	0.06
	(0.04)	(0.04)	(0.03)	(0.04)	(0.03)	(0.04)
num. children‡‡	0.01	-0.04*	0.01	-0.04*	0.01	-0.04**
	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)
grant		-0.40***		-0.41***		-0.40***
		(0.04)		(0.04)		(0.04)
household size		0.00		-0.00		0.00
		(0.01)		(0.01)		(0.01)
constant	6.87***	-3.26***	7.75***	-2.26***	7.54***	-2.61***
	(0.43)	(0.44)	(0.32)	(0.36)	(0.32)	(0.36)
N	5,808	5,808	6,460	6,460	6,436	6,436
F-stat	58.7	58.7	58.1	58.1	59.5	59.5
IMR	-0.61***		-0.60***		-0.60***	
	(0.06)		(0.05)		(0.05)	

Cluster-robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; † dummy variable (=1 if respondent is in the managerial/professional occupational category); †† dummy variable (=1 if respondent resides in a formal location; 0 otherwise); ‡ provincial unemployment rate; ‡‡ number of under-17 children in the household

Table A5.3: Heckman selection model of health-wage gradient in wave 3

	(1)	(2)	(3)	(4)	(5)	(6)
Variables	lnrealwage	employee	lnrealwage	employee	lnrealwage	employee
lnbmi	0.21***	0.30***				
	(0.07)	(0.09)				
1ncesd10			-0.03	-0.09**		
			(0.02)	(0.04)		
sah					0.15**	0.27***
					(0.06)	(0.07)
tenure	0.01***		0.01***		0.01***	
	(0.00)		(0.00)		(0.00)	
matric	0.40***	0.42***	0.40***	0.44***	0.40***	0.44***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
age	-0.02	0.07***	-0.01	0.07***	-0.01	0.07***
	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)
age^2	0.00	-0.00***	0.00	-0.00***	0.00	-0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
union	0.44***		0.46***		0.46***	
	(0.04)		(0.04)		(0.04)	
manprof†	0.48***		0.49***		0.50***	
	(0.04)		(0.04)		(0.04)	
tertiary	0.18***		0.19***		0.19***	
	(0.03)		(0.03)		(0.03)	
formal loc††	-0.11**	0.48***	-0.10**	0.47***	-0.10**	0.48***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
coloured	-0.12**	0.33***	-0.12**	0.29***	-0.12**	0.32***
	(0.05)	(0.07)	(0.05)	(0.07)	(0.05)	(0.07)
hours	0.01***		0.01***		0.01***	
	(0.00)		(0.00)		(0.00)	
male	0.29***	0.22***	0.26***	0.16***	0.26***	0.16***
	(0.03)	(0.04)	(0.03)	(0.04)	(0.03)	(0.04)
prov. unemp‡	0.01	-0.01*	0.01	-0.01*	0.01	-0.01*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
married	0.03	0.13***	0.04	0.14***	0.04	0.14***
	(0.03)	(0.04)	(0.03)	(0.04)	(0.03)	(0.04)
num. children‡‡	0.00	-0.01	0.00	-0.01	0.00	-0.02
	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)
grant		-0.35***		-0.34***		-0.34***
		(0.04)		(0.04)		(0.04)
household size		-0.01		-0.01		-0.01
		(0.01)		(0.01)		(0.01)
constant	6.98***	-2.42***	7.65***	-1.30***	7.41***	-1.77***
	(0.43)	(0.46)	(0.34)	(0.37)	(0.33)	(0.37)
N	5,089	5,089	5,186	5,186	5,184	5,184
F-stat	79.5	79.5	84.8	84.8	86.0	86.0
IMR	-0.70***		-0.70***		-0.69***	
	(0.06)		(0.06)		(0.06)	

Cluster-robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; † dummy variable (=1 if respondent is in the managerial/professional occupational category); †† dummy variable (=1 if respondent resides in a formal location; 0 otherwise); ‡ provincial unemployment rate; ‡‡ number of under-17 children in the household

Table A5.4: Censored quantile regression results of health-wage gradients across health definitions in wave 1 (negative zero coefficients due to rounding)

Variable	Q25	Q50	Q75	Q25	Q50	Q75	Q25	Q50	Q75
lnbmi	0.31***	0.33***	0.37***						
	(0.12)	(0.08)	(0.11)						
lncesd10				-0.07*	-0.06*	-0.06			
				(0.03)	(0.03)	(0.05)			
sah							0.22***	0.17**	0.20***
							(0.07)	(0.07)	(0.06)
tenure	0.02***	0.01***	0.01*	0.02***	0.01***	0.01	0.02***	0.01***	0.01
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
matric	0.60***	0.59***	0.61***	0.57***	0.59***	0.56***	0.61***	0.57***	0.59***
	(0.06)	(0.05)	(0.06)	(0.06)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
age	0.02	0.03*	0.02	0.03	0.03**	0.03	0.03	0.03**	0.02
	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
age^2	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
union	0.55***	0.58***	0.53***	0.59***	0.59***	0.55***	0.60***	0.61***	0.56***
	(0.06)	(0.06)	(0.06)	(0.06)	(0.05)	(0.05)	(0.06)	(0.06)	(0.05)
manprof†	0.35***	0.39***	0.41***	0.36***	0.40***	0.48***	0.32***	0.41***	0.47***
	(0.08)	(0.07)	(0.08)	(0.08)	(0.07)	(0.08)	(0.08)	(0.08)	(0.07)
tertiary	0.20***	0.16***	0.21***	0.17***	0.17***	0.24***	0.18***	0.16***	0.24***
	(0.07)	(0.06)	(0.06)	(0.06)	(0.05)	(0.05)	(0.07)	(0.06)	(0.05)
formal loc††	0.12*	0.16***	0.18***	0.12*	0.14***	0.18***	0.13*	0.16***	0.18***
	(0.07)	(0.05)	(0.06)	(0.07)	(0.05)	(0.06)	(0.07)	(0.06)	(0.06)
coloured	-0.00	0.02	0.04	-0.07	-0.02	-0.00	-0.06	-0.02	-0.02
	(0.06)	(0.07)	(0.08)	(0.06)	(0.06)	(0.07)	(0.06)	(0.06)	(0.07)
hours	0.01***	0.00**	0.00	0.00***	0.00**	0.00	0.00***	0.00**	0.00*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
male	0.30***	0.28***	0.36***	0.24***	0.23***	0.30***	0.25***	0.24***	0.30***
	(0.06)	(0.05)	(0.05)	(0.05)	(0.04)	(0.05)	(0.06)	(0.04)	(0.04)
prov. unemp‡	-0.01	-0.00	-0.01	-0.01	-0.00	-0.01	-0.02	-0.01	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
married	0.06	0.05	0.05	0.07	0.06	0.03	0.06	0.06	0.03
	(0.05)	(0.04)	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)	(0.05)	(0.05)
num. children‡‡	-0.03**	-0.03**	-0.02	-0.03*	-0.03**	-0.02	-0.03	-0.02**	-0.02
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)
constant	4.95***	5.14***	5.50***	6.20***	6.23***	6.66***	5.81***	6.02***	6.54***
	(0.61)	(0.41)	(0.46)	(0.36)	(0.33)	(0.41)	(0.34)	(0.33)	(0.36)
N	2189	2189	2189	2518	2518	2518	2505	2505	2505
Pseudo-R^2	0.26	0.33	0.36	0.27	0.34	0.36	0.27	0.34	0.36

Cluster-robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; †dummy variable (=1 if respondent is in the managerial/professional occupational category); ††dummy variable (=1 if respondent resides in a formal location; ‡provincial unemployment rate; ‡‡number of under-17 children in the household

Table A5.5: Censored quantile regression results of health-wage gradients across health definitions in wave 3 (negative zero coefficients due to rounding)

Variable	**Q25**	**Q50**	**Q75**	**Q25**	**Q50**	**Q75**	**Q25**	**Q50**	**Q75**
lnbmi	0.33***	0.41***	0.45***						
	(0.11)	(0.09)	(0.10)						
lncesd10				-0.06*	-0.07**	-0.05			
				(0.04)	(0.03)	(0.03)			
sah							0.22*	0.17**	0.22***
							(0.12)	(0.09)	(0.07)
tenure	0.01**	0.01**	0.01***	0.01***	0.01***	0.01***	0.01***	0.01***	0.01***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
matric	0.56***	0.54***	0.57***	0.55***	0.56***	0.58***	0.54***	0.57***	0.58***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
age	0.02	0.01	0.00	0.02	0.02	0.02	0.01	0.02	0.02
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
age^2	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
union	0.50***	0.49***	0.43***	0.51***	0.50***	0.45***	0.51***	0.51***	0.45***
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.07)	(0.06)	(0.06)	(0.06)
manprof†	0.53***	0.49***	0.46***	0.53***	0.48***	0.51***	0.53***	0.47***	0.50***
	(0.06)	(0.06)	(0.07)	(0.06)	(0.06)	(0.07)	(0.06)	(0.06)	(0.07)
tertiary	0.19***	0.24***	0.25***	0.19***	0.26***	0.24***	0.18***	0.25***	0.23***
	(0.06)	(0.05)	(0.05)	(0.06)	(0.05)	(0.05)	(0.06)	(0.05)	(0.05)
formal loc††	0.22***	0.10**	0.09*	0.19**	0.10**	0.09**	0.21***	0.13***	0.10**
	(0.08)	(0.05)	(0.05)	(0.08)	(0.05)	(0.05)	(0.08)	(0.05)	(0.05)
coloured	0.03	0.02	0.01	-0.02	-0.01	-0.00	-0.01	0.02	-0.00
	(0.07)	(0.05)	(0.06)	(0.07)	(0.05)	(0.07)	(0.07)	(0.05)	(0.07)
hours	0.01***	0.00	0.00**	0.01***	0.00*	0.00*	0.01***	0.00*	0.00*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
male	0.34***	0.40***	0.41***	0.25***	0.30***	0.35***	0.25***	0.31***	0.35***
	(0.05)	(0.04)	(0.05)	(0.05)	(0.04)	(0.05)	(0.05)	(0.04)	(0.05)
prov. unemp‡	-0.00	0.00	0.00	-0.00	0.00	0.01	-0.00	-0.00	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
married	0.04	0.04	0.10**	0.08	0.03	0.12**	0.07	0.05	0.12***
	(0.06)	(0.04)	(0.04)	(0.06)	(0.04)	(0.05)	(0.05)	(0.04)	(0.04)
num. children‡‡	-0.04*	-0.03*	-0.02	-0.04**	-0.02	-0.01	-0.03**	-0.02	-0.01
	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)
constant	4.92***	5.32***	5.44***	6.13***	6.50***	6.69***	5.83***	6.29***	6.48***
	(0.59)	(0.48)	(0.46)	(0.48)	(0.41)	(0.40)	(0.49)	(0.46)	(0.37)
N	2747	2747	2747	2806	2806	2806	2805	2805	2805
Pseudo R^2	0.26	0.33	0.36	0.26	0.33	0.35	0.26	0.33	0.36

Cluster-robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; †dummy variable (=1 if respondent is in the managerial/professional occupational category); ††dummy variable (=1 if respondent resides in a formal location); ‡provincial unemployment rate; ‡‡number of under-17 children in the household

Table A5.6: Gender-based censored quantile regression results of health-wage gradients across health definitions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
			lnbmi						lncesd10						sah			
	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M
	Q25	Q25	Q50	Q50	Q75	Q75	Q25	Q25	Q50	Q50	Q75	Q75	Q25	Q25	Q50	Q50	Q75	Q75
									Wave 1									
	0.19	0.47**	0.14	0.56***	0.16	0.69***	-0.08*	-0.03	-0.07	-0.06	-0.09	-0.05	0.30**	0.14	0.18**	0.14*	0.18***	0.18*
	(0.13)	(0.17)	(0.12)	(0.13)	(0.13)	(0.19)	(0.05)	(0.05)	(0.05)	(0.04)	(0.05)	(0.07)	(0.10)	(0.11)	(0.08)	(0.08)	(0.06)	(0.10)
									Wave 3									
	0.17	0.62***	0.12	0.76***	0.20*	0.82***	-0.04	-0.09**	-0.06	-0.06*	-0.09**	-0.01	0.29**	0.20*	0.15	0.24**	0.20***	0.25*
	(0.16)	(0.18)	(0.10)	(0.16)	(0.12)	(0.18)	(0.05)	(0.05)	(0.04)	(0.04)	(0.03)	(0.04)	(0.14)	(0.11)	(0.12)	(0.11)	(0.07)	(0.14)
N(W1†)	1065	1124	1065	1124	1065	1124	1221	1297	1221	1297	1221	1297	1215	1290	1215	1290	1215	1290
R^2(W1)	0.28	0.23	0.35	0.30	0.43	0.30	0.28	0.23	0.36	0.30	0.43	0.29	0.29	0.24	0.36	0.30	0.43	0.29
N(W3††)	1433	1314	1433	1314	1433	1314	1464	1342	1464	1342	1464	1342	1463	1342	1463	1342	1463	1342
R^2(W3)	0.30	0.21	0.36	0.28	0.42	0.29	0.30	0.21	0.37	0.25	0.43	0.27	0.30	0.21	0.37	0.25	0.43	0.28
controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Cluster-robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1; F ⇒ females; M ⇒males; R^2 ⇒ pseudo R^2; controls in each regression are as in the main regression; † Wave 1; †† Wave 3

Table A6.1: Impairment-related Blinder-Oaxaca log wage decomposition (memory loss and tuberculosis only)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	
Non-discriminatory norm	N/NG1/NG2	Raw gap	Xteristics†	Returns	Selection	Estimated gap	Percentage contribution to estimated gap	
							Xteristics	Returns
Wave 1								
OLS estimates	10,273/2403/117	0.44	0.32***	0.10		0.42***	76.2	23.8
			(0.09)	(0.08)		(0.11)		
Selection-corrected estimates	11,213/2403/117	0.44	0.25***	0.80***	-0.82*	1.05***	23.8	76.2
			(0.07)	(0.17)	(0.48)	(0.19)		
Wave 3								
OLS estimates	7,871/2668/138	0.43	0.24***	0.23***		0.48***	50.0	47.8
			(0.06)	(0.07)		(0.08)		
Selection-corrected estimates	8,661/2668/138	0.43	0.19***	0.71***	-0.53	0.90***	21.1	78.9
			(0.05)	(0.26)	(0.51)	(0.26)		

Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1; statistics corrected for complex survey design and national representativeness; standard errors of difference between selection terms obtained via bootstrapping with 400 replications; non-impaired wage profile serves as non-discriminatory norm; N⟹ sample size; N1⟹ number of observations in non-impaired group; N2⟹ number of observations in impaired group; †characteristics/endowment

Table A6.2: Ordinary least squares log wage regressions

	(1)	(2)	(3)	(4)
	Wave 1		Wave 3	
Variables	Non-impaired	Impaired	Non-impaired	Impaired
schooling	0.09***	0.08***	0.10***	0.03
	(0.01)	(0.02)	(0.01)	(0.02)
age	0.04**	-0.01	0.03*	-0.02
	(0.02)	(0.06)	(0.02)	(0.06)
age^2	-0.00*	0.00	-0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)
urban formal location	0.30***	0.31**	0.20***	0.36**
	(0.05)	(0.14)	(0.04)	(0.18)
coloured	0.09	0.32	0.02	-0.35
	(0.07)	(0.30)	(0.08)	(0.25)
male	0.34***	0.18	0.31***	0.12
	(0.04)	(0.12)	(0.04)	(0.13)
Num. of U-17 children†	-0.03**	-0.06*	-0.02	-0.08*
	(0.01)	(0.03)	(0.01)	(0.05)
provincial unemployment rate	-0.00	0.03	-0.00	-0.03
	(0.01)	(0.03)	(0.01)	(0.02)
married/cohabiting	0.08*	0.37***	0.13***	0.23
	(0.04)	(0.12)	(0.04)	(0.15)
tenure	0.01	0.02***	0.02***	0.02***
	(0.00)	(0.01)	(0.00)	(0.01)
union	0.40***	0.43***	0.33***	0.14
	(0.05)	(0.14)	(0.05)	(0.14)
managerial/professional	0.41***	0.36*	0.49***	0.51**
	(0.07)	(0.22)	(0.05)	(0.24)
tertiary	0.08	0.12	0.13***	0.18
	(0.06)	(0.17)	(0.05)	(0.16)
hours	0.00	-0.00	0.00	0.01
	(0.00)	(0.00)	(0.00)	(0.00)
cesd10	-0.02***	0.00	-0.01*	-0.03
	(0.00)	(0.01)	(0.01)	(0.02)
constant	5.48***	5.63***	5.65***	8.03***
	(0.35)	(1.11)	(0.38)	(1.36)
R^2	0.50	0.63	0.48	0.39
N	2379	139	2661	145
F-stat	70.2	13.0	86.3	4.8

Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1; estimates corrected for complex survey design, national representativeness and non-random attrition; † Number of under-17 children in the household

Table A6.3: Selection-corrected log wage regressions (wave 1)

	(1)	(2)	(3)	(4)
	Non-impaired		Impaired	
Variables	lnrwage†	employee	lnrwage	employee
schooling	0.07***	0.05***	0.08***	0.04
	(0.01)	(0.01)	(0.02)	(0.03)
age	-0.01	0.11***	0.00	0.26***
	(0.02)	(0.02)	(0.06)	(0.07)
age^2	0.00	-0.00***	0.00	-0.00***
	(0.00)	(0.00)	(0.00)	(0.00)
urban formal location	0.25***	0.11	0.34**	0.49**
	(0.06)	(0.08)	(0.15)	(0.19)
coloured	0.03	0.14	0.31	-0.17
	(0.09)	(0.13)	(0.30)	(0.24)
male	0.21***	0.32***	0.19	0.03
	(0.05)	(0.05)	(0.12)	(0.21)
number of under-17 children in household	0.02*	-0.02	-0.06*	-0.19**
	(0.01)	(0.03)	(0.03)	(0.08)
provincial unemployment rate	-0.00	-0.01	0.02	-0.03
	(0.01)	(0.02)	(0.03)	(0.04)
married/cohabiting	0.04	0.09	0.37***	-0.02
	(0.04)	(0.06)	(0.12)	(0.20)
tenure	0.01		0.02***	
	(0.00)		(0.01)	
union	0.39***		0.43***	
	(0.05)		(0.14)	
managerial/professional	0.42***		0.36	
	(0.07)		(0.22)	
tertiary	0.08		0.12	
	(0.06)		(0.16)	
hours	0.00		-0.00	
	(0.00)		(0.00)	
cesd10	-0.01	-0.03***	0.00	-0.04**
	(0.01)	(0.01)	(0.01)	(0.02)
grant		-0.45***		-0.39*
		(0.06)		(0.22)
household size		-0.02		0.08**
		(0.02)		(0.04)
constant	7.21***	-2.42***	5.43***	-4.55***
	(0.43)	(0.50)	(1.23)	(1.35)
N	6019	6019	441	441
F-stat	31.7	31.7	10.9	10.9
IMR	-0.62***		0.07	
	(0.05)		(0.18)	

Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; estimates corrected for complex survey design, national representativeness and non-random attrition; † Natural logarithm of real monthly wage

Table A6.4: Selection-corrected log wage regressions (wave 3)

	(1)	(2)	(3)	(4)
	Non-impaired		Impaired	
Variables	lnrwage†	employee	lnrwage	employee
schooling	0.07***	0.06***	0.03	0.08***
	(0.01)	(0.01)	(0.03)	(0.03)
age	-0.01	0.09***	-0.02	-0.08
	(0.02)	(0.03)	(0.06)	(0.09)
age^2	0.00	-0.00***	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)
urban formal location	0.06	0.23***	0.36**	-0.13
	(0.05)	(0.07)	(0.18)	(0.21)
coloured	-0.10	0.33**	-0.35	-0.34
	(0.11)	(0.15)	(0.27)	(0.35)
male	0.23***	0.16***	0.11	0.20
	(0.05)	(0.06)	(0.14)	(0.21)
number of under-17 children in household	0.02	-0.06**	-0.08*	0.17*
	(0.01)	(0.03)	(0.05)	(0.09)
provincial unemployment rate	0.00	-0.01	-0.03	0.01
	(0.01)	(0.01)	(0.02)	(0.03)
married/cohabiting	0.06	0.25***	0.23	-0.18
	(0.04)	(0.07)	(0.15)	(0.22)
tenure	0.02***		0.02***	
	(0.00)		(0.01)	
union	0.32***		0.14	
	(0.05)		(0.14)	
managerial/professional	0.48***		0.51**	
	(0.05)		(0.24)	
tertiary	0.11***		0.18	
	(0.04)		(0.16)	
hours	0.00		0.01	
	(0.00)		(0.00)	
cesd10	-0.00	-0.01	-0.03	-0.03
	(0.01)	(0.01)	(0.02)	(0.02)
grant		-0.44***		-0.28
		(0.06)		(0.21)
household size		0.00		-0.11**
		(0.01)		(0.05)
constant	7.14***	-2.21***	8.04***	0.39
	(0.51)	(0.58)	(1.41)	(1.98)
N	4840	4840	345	345
F-stat	42.88	42.88	4.315	4.315
IMR	-0.68***		-0.01	
	(0.07)		(0.31)	

Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; estimates corrected for complex survey design, national representativeness and non-random attrition; † Natural logarithm of real monthly wage

Table A6.5: Wage gap components across deciles of the wage distribution (t values in parentheses): wave 1 and wave 3

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Wave 1				Wave 3			
Quantiles	total	characteristics	returns	residual	total	characteristics	returns	residual
0.1	0.58	0.33	0.28	-0.03	0.31	0.20	0.17	-0.06
	(2.71)	(3.58)	(1.84)	(-0.20)	(1.60)	(2.56)	(0.89)	(-0.54)
0.2	0.48	0.28	0.23	-0.03	0.31	0.18	0.19	-0.06
	(2.64)	(3.27)	(1.91)	(-0.28)	(2.56)	(2.42)	(1.41)	(-0.74)
0.3	0.42	0.25	0.19	-0.02	0.33	0.19	0.19	-0.06
	(2.75)	(3.01)	(1.76)	(-0.24)	(3.04)	(2.82)	(1.82)	(-0.76)
0.4	0.41	0.25	0.17	-0.01	0.34	0.22	0.20	-0.08
	(2.70)	(2.97)	(1.68)	(-0.10)	(4.52)	(3.46)	(2.12)	(-1.32)
0.5	0.44	0.25	0.16	0.03	0.48	0.24	0.20	0.04
	(2.88)	(3.03)	(1.62)	(0.39)	(6.08)	(3.70)	(2.05)	(0.52)
0.6	0.47	0.27	0.14	0.06	0.51	0.24	0.21	0.05
	(3.01)	(3.34)	(1.47)	(0.69)	(7.60)	(3.69)	(2.13)	(0.80)
0.7	0.46	0.31	0.11	0.04	0.58	0.26	0.21	0.11
	(3.84)	(4.10)	(1.11)	(0.44)	(6.88)	(3.78)	(1.97)	(1.54)
0.8	0.47	0.34	0.09	0.03	0.62	0.29	0.21	0.12
	(3.59)	(4.35)	(0.91)	(0.36)	(5.87)	(4.14)	(1.72)	(1.48)
0.9	0.56	0.37	0.05	0.14	0.67	0.32	0.24	0.11
	(4.28)	(4.53)	(0.42)	(1.29)	(6.09)	(3.90)	(1.68)	(1.05)

Select health questions as they occur in the questionnaire

SAH was derived from the following question:

Question: How would you describe your health at present? Would you say it is excellent, very good, good, fair or poor?

Possible responses: Excellent [1]; Very good [2]; Good [3]; Fair [4]; Poor [5]

BMI was obtained from height and weight questions. It starts with the following general question: Now, we would like to take your height, weight, waist and blood pressure measurements.

Respondent's height (centimetres)

Respondent's weight (kilograms)

REFERENCES

Abadie, A. (2000b). Semiparametric estimation of instrumental variable models for causal effects. *Technical Working Paper 260, National Bureau of Economic Research, Cambridge, MA.*

Acemoglu, D., & Angrist, J. D. (2001). Consequences of employment protection? The case of the Americans with Disabilities Act. *Journal of Political Economy, 109*(5), 915-957.

African Economic Outlook. (2012). Education and skills mismatch, accessed 1 June, 2015, <http://www.africaneconomicoutlook.org/en/theme/youth_employment/education-skills-mismatch/>

African Economic Outlook. (2014). *African economic outlook 2014: South Africa.* Tunis: African Development Bank.

Aigner, D. J., & Cain, G. G. (1977). Statistical theories of discrimination in labor markets. *Industrial and Labor Relations Review, 30*(2), 175-187.

Alderman, H., Chiappori, P., Haddad, L., Hoddinott, J., & Kanbur, R. (1995). Unitary versus collective models of the household: Is it time to shift the burden of proof? *The World Bank Research Observer, 10*(1), 1-19.

Altonji, J. G., & Blank, R. M. (1999). Race and gender in the labor market. In O. Ashelfelter, & D. Card (Eds.), *Handbook of labor economics* (3rd ed., pp. 3143-3259). Amsterdam: Elsevier.

Anderson, K. H., & Burkhauser, R. V. (1984). The importance of the measure of health in empirical estimates of the labor supply of older men. *Economics Letters, 16*(3-4), 375-380.

Anderson, K. H., & Burkhauser, R. V. (1985). The retirement-health nexus: A new measure of an old puzzle. *Journal of Human Resources, 20*(3), 315-330.

Anderson, P. M., & Levine, P. B. (1999). Child care and mothers' employment decisions. *NBER Working Paper no. 7058.*

Andrén, D., & Palmer, E. (2001). The effect of sickness on earnings. *Göteborg University, School of Business, Economics and Law, Working Paper no. 45,* 37-73.

Andresen, E. M., Malmgren, J. A., Carter, W. B., & Patrick, D. L. (1994). Screening for depression in well older adults: Evaluation of a short form of the CES-D. *American Journal of Preventive Medicine, 10*(2), 77-84.

Angrist, J. D. (2001). Estimation of limited dependent variable models with dummy endogenous regressors. *Journal of Business & Economic Statistics, 19*(1), 1-28.

Angrist, J. D., & Pischke, J. (2009). *Mostly harmless econometrics: An empiricist's companion.* Princeton, New Jersey: Princeton University Press.

Antonczyk, D., Fitzenberger, B., & Sommerfeld, K. (2010). Rising wage inequality, the decline of collective bargaining, and the gender wage gap. *Labour Economics, 17*(5), 835-847.

Ardington, C., & Case, A. (2009). Health: Analysis of the NIDS wave 1 dataset. *NIDS Discussion Paper, no. 2.*

Ardington, C., Case, A., & Hosegood, V. (2009). Labor supply responses to large social transfers: Longitudinal evidence from South Africa. *American Economic Journal: Applied Economics, 1*(1), 22-48.

Ardington, C., & Gasealahwe, B. (2014). Mortality in South Africa: Socio-economic profile and association with self-reported health. *Development Southern Africa, 31*(1), 127-145.

Arndt, C., & Lewis, J. D. (2001). The HIV/AIDS pandemic in South Africa: Sectoral impacts and unemployment. *Journal of International Development, 13*(4), 427-449.

Arrow, K. (1973). The theory of discrimination. In O. Ashenfelter, & A. Rees (Eds.), *Discrimination in the labour market* (pp. 3-33). Princeton: Princeton University Press.

Ataguba, J. E. (2012). *Distributional impact of health care finance in South Africa.* (Unpublished Ph.D.). University of Cape Town, Cape Town.

Au, D., Crossley, T. F., & Schellhorn, M. (2005). The effect of health changes and long-term health on the work activity of older Canadians. *Health Economics, 14*(10), 999-1018.

AVERT (undated). South Africa HIV/AIDS statistics, accessed on 30/05/2015, <http://www.avert.org/south-Africa-hiv-aids-statistics.htm>

Baldwin, M., & Johnson, W. G. (1994). Labor market discrimination against men with disabilities. *Journal of Human Resources, 29*(1), 1-19.

Baldwin, M. L., & Johnson, W. G. (2006). A critical review of studies of discrimination against workers with disabilities. In W. M. Rodgers III (Ed.), *Handbook on the economics of discrimination* (pp. 119-160). Cheltenham, UK: Edward Elgar.

Bancroft, G. (1958). *The American labor force: Its growth and changing composition.* New York: Wiley.

Banerjee, A., Galiani, S. F., Levinsohn, J. A., McLaren, Z., & Woolard, I. (2008). Why has unemployment risen in the new South Africa. *Economics of Transition, 16*(4), 715-740.

Bartel, A., & Taubman, P. (1986). Some economic and demographic consequences of mental illness. *Journal of Labor Economics, 4*, 243-256.

Basta, S. S., Karyadi, D., & Scrimshaw, N. S. (1979). Iron deficiency anemia and the productivity of adult males in indonesia. *American Journal of Clinical Nutrition, 32*(4), 916-925.

Baum, C. L., & Ford, W. F. (2004). The wage effects of obesity: A longitudinal study. *Health Economics, 13*(9), 885-899.

Bazzoli, G. J. (1985). The early retirement decision: New empirical evidence on the influence of health. *Journal of Human Resources,* , 214-234.

Beblo, M., Beninger, D., Heinze, A., & Laisney, F. (2003). Measuring selectivity-corrected gender wage gaps in the EU. *ZEW Discussion Paper no. 03-74.*

Becker, G. S. (1964). *Human capital.* New York: Columbia University Press.

Becker, G. S. (1981). *A treatise on the family* Cambridge, Harvard University Press.

Becker, G. S. (1965). A theory of the allocation of time. *The Economic Journal, 75*(299), 493-517.

Becker, G. S. (1971). *The economics of discrimination* (2nd ed.). Chicago: University of Chicago Press.

Becker, G. S. (1973). A theory of marriage: Part I. *The Journal of Political Economy, 81*(4), 813-846.

Beckett, M., & Elliott, M. N. (2002). Does the association between marital status and health vary by sex, race, and ethnicity? *Rand Working Paper Series 02-08.*

Behrman, J. R. (2009). Early life nutrition and subsequent education, health, wage, and intergenerational effects. In M. Spence, & M. Lewis (Eds.), *Health and growth* (pp. 167-183) World Bank Publications.

Behrman, J. R., & Rosenzweig, M. R. (2004). Returns to birthweight. *Review of Economics and Statistics, 86*(2), 586-601.

Benítez-Silva, H., Buchinsky, M., Chan, H. M., Cheidvasser, S., & Rust, J. (2004). How large is the bias in self-reported disability? *J.Appl.Econ, 19*, 649-670.

Bertrand, M., Mullainathan, S., & Miller, D. (2003). Public policy and extended families: Evidence from pensions in South Africa. *The World Bank Economic Review, 17*(1), 27-50.

Bhorat, H. (2004). Labour market challenges in the post-apartheid South Africa. *South African Journal of Economics, 72*(5), 940-977.

Bhorat, H. (2007). Unemployment in South Africa: Descriptors and determinants. *Unpublished Consultant Report for OECD, Paris.*

Bhorat, H., & Hodge, J. (1999). Decomposing shifts in labour demand in South Africa. *South African Journal of Economics, 67*(3), 155-168.

Bhorat, H., & Leibbrandt, M. (1999b). Modelling vulnerability and low earnings for South African labour market.

Blanchflower, D. G., & Oswald, A., J. (1994). *The wage curve* MIT press.

Blinder, A. S. (1973). Wage discrimination: Reduced form and structural estimates. *Journal of Human Resources, 8*(4), 436-455.

Bloom, D. E., Canning, D., & Graham, B. (2002). Longevity and life cycle savings. *NBER Working Paper 8808.*

Blundell, R., & MaCurdy, T. (1999). Labor supply: A review of alternative approaches. In O. Ashenfelter, & D. Card (Eds.), *Handbook of labour economics* (Vol 3A ed., pp. 1559-1695). Amsterdam: Elsevier.

Booysen, F. R., Bachmann, M., Matebesi, Z., & Meyer, J. (2002). The socioeconomic impact of HIV/AIDS on households in South Africa: Pilot study in welkom and qwaqwa, free state province. *University of the Free State.*

Boskin, M. J. (1977). Social security and retirement decisions. *Economic Inquiry, 15*(1), 1-25.

Bound, J. (1989). The health and earnings of rejected disability insurance applicants. *NBER Working Paper no. 2816.*

Bound, J. (1991). Self-reported versus objective measures of health in retirement models. *Journal of Human Resources,* (1), 138.

Bound, J., Schoenbaum, M., Stinebrickner, T., & Waidmann, T. (1999). The dynamic effects of health on the labor force transitions of older workers. *Labour Economics, 6*(2), 179-202.

Bound, J., Schoenbaum, M., & Waidmann, T. (1995). Race and education differences in disability status and labour force attachment in the health and retirement survey. *Journal of Human Resources, 30,* S227-S267.

Bowen, W. G., & Finegan, T. A. (1969). *The economics of labor force participation* Princeton University Press Princeton.

Boymond, M., Flückiger, Y., & Silber, J. (1994). Wage discrimination and occupational segregation by gender: Some evidence from swiss data. *Départment d'Economie Politique, Université De Genèva, Mimeo.*

Bradshaw, D., Groenewald, P., Laubscher, R., Nannan, N., Nojilana, B., Norman, R., . . . Schneider, M. (2000). *Initial burden of disease estimates for South Africa.* Cape Town: South African Medical Research Council.

Bridges, S., & Lawson, D. (2008). Health and labour market participation in Uganda. *United Nations University World Institute for Development Economics Research Discussion Paper, No 2008/07.*

Brunello, G., & d'Hombres, B. (2007). Does body weight affect wages?: Evidence from europe. *Economics & Human Biology, 5*(1), 1-19.

Butcher, K. F., & Rouse, C. E. (2001). Wage effects of unions and industrial councils in South Africa. *Indus.& Lab.Rel.Rev., 54*(2), 349-374.

Cahuc, P., & Zylberberg, A. (2004). *Labor economics* MIT Press.

Cai, L. (2010). The relationship between health and labour force participation: Evidence from a panel data simultaneous equation model. *Labour Economics, 17*(1), 77-90.

Cai, L., & Kalb, G. (2006). Health status and labour force participation: Evidence from australia. *Health Economics, 15*(3), 241-261.

Cameron, A. C., & Trivedi, P. K. (2005). *Microeconometrics: Methods and applications*. New York: Cambridge University Press.

Cameron, A. C., & Trivedi, P. K. (2010). *Microeconometrics using stata* Stata Press College Station, TX.

Campolieti, M. (2002). Disability and the labor force participation of older men in canada. *Labour Economics, 9*(3), 405-432.

Casale, D. (2004). What has the feminisation of the labour market 'bought' women in South Africa?: Trends in labour force participation, employment and earnings, 1995-2001. *Journal of Interdisciplinary Economics, 15*(3/4), 251-276.

Casale, D., & Posel, D. (2002). The continued feminisation of the labour force in South Africa. *South African Journal of Economics, 70*(1), 156-184.

Case, A., Paxson, C., & Ableidinger, J. (2004). Orphans in Africa: Parental death, poverty, and school enrollment. *Demography, 41*(3), 483-508.

Cawley, J. (2004). The impact of obesity on wages. *Journal of Human Resources, 39*(2), 451-474.

Chamberlain, D., & Van der Berg, S. (2002). Earnings functions, labour market discrimination and quality of education in South Africa.

Chiappori, P. (1992). Collective labor supply and welfare. *Journal of Political Economy, 100*(3), 437-467.

Chiburis, R. C., Das, J., & Lokshin, M. (2011). A practical comparison of the bivariate probit and linear IV estimators. *Policy Research Working Paper 5601.*

Chirikos, T. N. (1993). The relationship between health and labor market status. *Annual Review of Public Health, 14*(1), 293-312.

Chirikos, T. N., & Nestel, G. (1985). Further evidence on the economic effects of poor health. *The Review of Economics and Statistics,* , 61-69.

Choi, H. (2007). *Health impacts on future labor market outcome: Exploring PSID siblings.* (Unpublished).

Cichello, P., Leibbrandt, M., & Woolard, I. (2012). Labour market: Analysis of the NIDS wave 1 and 2 datasets. *(SALDRU Working Paper Number 78 / NIDS Discussion Paper 2012/1).*

Coate, S., & Loury, G. C. (1993). Will affirmative-action policies eliminate negative stereotypes? *The American Economic Review, 83*(5), 1220-1240.

Contoyannis, P., & Rice, N. (2001). The impact of health on wages: Evidence from the british household panel survey. *Empirical Economics, 26*(4), 599-622.

Cortez, R. (1999). *Salud y productividad en el perú: Un análisis empírico por género y región,* Banco InterAmericano de Desarrollo.

Croppenstedt, A., & Muller, C. (2000). The impact of farmers' health and nutritional status on their productivity and efficiency: Evidence from Ethiopia. *Economic Development and Cultural Change, 48*(3), 475-502.

Currie, J. (1995). Socio-economic status and child health: Does public health insurance narrow the gap? *The Scandinavian Journal of Economics, 97*(4), 603-620.

Currie, J., & Madrian, B. C. (1999). Health, health insurance and the labor market. *Handbook of Labor Economics, 3*, 3309-3416.

Deaton, A. (2002). Policy implications of the gradient of health and wealth. *Health Affairs, 21*(2), 13-30.

DeLeire, T. (2001). Changes in wage discrimination against people with disabilities: 1984-93. *Journal of Human Resources, 36*(1), 144-158.

Department of Health, Medical Research Council, & OrcMacro. (2007). *South Africa demographic and health survey, 2003.* Pretoria: Department of Health.

DeSalvo, K. B., Bloser, N., Reynolds, K., He, J., & Muntner, P. (2006). Mortality prediction with a single general Self-Rated health question. *Journal of General Internal Medicine, 21*(3), 267-275.

Development Indicators. (2010). *Development indicators, 2010.* Pretoria: The Presidency, Republic of South Africa.

Dinkelman, T., & Pirouz, F. (2002). Individual, household and regional determinants of labour force attachment in South Africa. *South African Journal of Economics, 70*(5), 865-891.

Douglas, P. H. (1934). *The theory of wages*. New York: MacMillan Company.

Dow, W. H., Gertler, P., Schoeni, R. F., Strauss, J., & Thomas, D. (1997). *Health care prices, health and labor outcomes: Experimental evidence. (*Unpublished manuscript).

du Toit, J. (2002). The structure of the South African economy. *The SA Financial Forum.*

Duflo, E. (2003). Grandmothers and granddaughters: Old-age pensions and intrahousehold allocation in South Africa. *World Bank Economic Review, 17*(1), 1-25.

Dwyer, D. S., & Mitchell, O. S. (1999). Health problems as determinants of retirement: Are self-rated measures endogenous? *Journal of Health Economics, 18*(2), 173-193.

Ettner, S. L., Frank, R. G., & Kessler, R. C. (1997). The impact of psychiatric disorders on labor market outcomes. *Industrial and Labour Relations Review,* , 81.

Ettner, S. L. (2000). The relationship between labor market outcomes and physical and mental health exogenous human capital or endogenous health production? In D. S. Salkever, & A. Sorkin (Eds.), *The economics of disability* (pp. 1-31) Emerald Group Publishing Limited.

Evans, P., & McCormick, B. (1994). The new pattern of regional unemployment: Causes and policy significance. *The Economic Journal, 104*(424), 633-647.

Ferraro, K. F. (1980). Self-ratings of health among the old and the old-old. *Journal of Health and Social Behavior,* , 377-383.

Field, E., Robles, O., & Torero, M. (2009). Iodine deficiency and schooling attainment in tanzania. *American Economic Journal: Applied Economics, 1*(4), 140-169.

Fogel, R. W. (1994). Economic growth, population theory, and physiology: The bearing of long-term processes on the making of economic policy. *American Economic Review, 84*(3), 369-395.

Fortin, B., & Lacroix, G. (1997). A test of the unitary and collective models of household labour supply. *The Economic Journal, 107*(443), 933-955.

Fortin, N., Lemieux, T., & Firpo, S. (2011). Decomposition methods in economics. In O. Ashenfelter, & D. Card (Eds.), *Handbook of labour economics* (4A ed., pp. 1-102) Elsevier.

Foster, A. D., & Rosenzweig, M. R. (1992). Information flows and discrimination in labor-markets in rural-areas in developing-countries. *World Bank Economic Review,* , 173-203.

Foster, A. D., & Rosenzweig, M. R. (1994). A test for moral hazard in the labor market: Contractual arrangements, effort, and health. *The Review of Economics and Statistics, 7*(2), 213-227.

Frankenberg, E., & Jones, N. R. (2004). Self-rated health and mortality: Does the relationship extend to a low income setting? *Journal of Health and Social Behavior, 45*(4), 441-452.

Freeman, R. B. (1982). Union wage practices and wage dispersion within establishments. *Industrial and Labor Relations Review, 36*(1), 3-21.

Fuchs, V. R. (1972). The contribution of health services to the American economy. *National Bureau of Economic Research,* , 1-38.

Godlonton, S., & Keswell, M. (2005). The impact of health on poverty: Evidence from the South African integrated family survey. *South African Journal of Economics, 73*(1), 133-148.

Goetzel, R. Z., Long, S. R., Ozminkowski, R. J., Hawkins, K., Wang, S., & Lynch, W. (2004). Health, absence, disability, and presenteeism cost estimates of certain physical and mental health conditions affecting US employers. *Journal of Occupational and Environmental Medicine, 46*(4), 398-412.

Gómez, P. G., & Nicolás, A. L. (2006). Health shocks, employment and income in the spanish labour market. *Health Economics, 15*(9), 997-1009.

Griliches, Z. (1977). Estimating the returns to schooling: Some econometric problems. *Econometrica: Journal of the Econometric Society, 45*, 1-22.

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

Grün, C. (2004). Direct and indirect gender discrimination in the South African labour market. *International Journal of Manpower, 25*(3/4), 321-342.

Haan, P., & Myck, M. (2009). *Dynamics of poor health and non-employment* Forschungsinst. zur Zukunft der Arbeit.

Haveman, R., Wolfe, B., Kreider, B., & Stone, M. (1994). Market work, wages, and men's health. *Journal of Health Economics, 13*(2), 163-182.

Haveman, R., Wolfe, B. L., & Huang, F. M. (1989). Disability status as an unobservable: Estimates from a structural model. *NBER Working Paper no. 2831.*

Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *NBER Working Paper no. 12006.*

Heckman, J. J., Killingsworth, M., & MaCurdy, T. (1981). Empirical evidence on static labour supply models: A survey of recent developments. In Z. Hornstein, J. Grice & A. Webb (Eds.), *The economics of the labour market* (pp. 75-122). London: Her Majesty's Stationery Office.

Heckman, J. J., & MaCurdy, T. E. (1981). New methods for estimating labor supply functions: A survey. *Research in Labor Economics, 4*, 65-102.

Hicks, J. R. (1946). *Value and capital: An inquiry into some fundamental principles of economic theory* (2nd ed.). Oxford: Clarendon Press.

Hlekiso, T., & Mahlo, N. (2006). Wage trends and inequality in South Africa: A comparative analysis. *South African Reserve Bank, Labour Market Frontiers, October,* (8), 9-16.

Hoogeveen, J. G. (2005). Measuring welfare for small but vulnerable groups: Poverty and disability in uganda. *Journal of African Economies, 14*(4), 603-631.

Husain, M. J. (2010). Contribution of health to economic development: A survey and overview. *Economics: The Open Access, Open-Assessment E-Journal, 4.*

Ilkkaracan, I., & Selim, R. (2003). The role of unemployment in wage determination: Further evidence on the wage curve from turkey. *Applied Economics, 35*(14), 1589-1598.

Jack, W., & Lewis, M. (2009). Health investments and economic growth: Macroeconomic evidence and microeconomic foundations. *World Bank Policy Research Working Paper no.4877.*

Jäckle, R., & Himmler, O. (2010). Health and wages-panel data estimates considering selection and endogeneity. *Journal of Human Resources, 45*(2), 364-406.

Jamison, D. T., Summers, L. H., Alleyne, G., Arrow, K. J., Berkley, S., Binagwaho, A., . . . Frenk, J. (2013). Global health 2035: A world converging within a generation. *The Lancet, 382*(9908), 1898-1955.

Jann, B. (2008). The Blinder-Oaxaca decomposition for linear regression models. *The Stata Journal, 8*(4), 453-479.

Jaumotte, F. (2003). Female labour force participation: Past trends and main determinants in OECD countries. *OECD Economics Department Working Papers, no. 376.*

Johnson, W. G., & Lambrinos, J. (1985). Wage discrimination against handicapped men and women. *Journal of Human Resources, 20*, 264-277.

Joll, C., McKenna, C. J., McNabb, R., & Shorey, J. (1983). *Developments in labour market analysis* G. Allen & Unwin.

Jones, M. K., Latreille, P. L., & Sloane, P. J. (2006). Disability, gender, and the british labour market. *Oxford Economic Papers, 58*(3), 407-449.

Jones, M. K. (2006). Is there employment discrimination against the disabled? *Economics Letters, 92*(1), 32-37.

Juhn, C., Murphy, K. M., & Pierce, B. (1993). Wage inequality and the rise in returns to skill. *Journal of Political Economy, 101*(3), 410-442.

Jylhä, M. (2009). What is self-rated health and why does it predict mortality? towards a unified conceptual model. *Social Science & Medicine, 69*(3), 307-316.

Kenkel, D. S. (1995). Should you eat breakfast? estimates from health production functions. *Health Economics, 4*(1), 15-29.

Kessler, R., Heeringa, S., Lakoma, M., Petukhova, M., Rupp, A., Schoenbaum, M., . . . Zaslavsky, A. (2008). Individual and societal effects of mental disorders on earnings in the united states: Results from the national comorbidity survey replication. *American Journal of Psychiatry, 165*(6), 703-711.

Kidd, M. P., Sloane, P. J., & Ferko, I. (2000). Disability and the labour market: An analysis of British males. *Journal of Health Economics, 19*(6), 961-981.

Killingsworth, M. (1981). A survey of labor supply models: Theoretical analysis and first-generation empirical results. *Research in Labor Economics, 4*, 1-64.

Killingsworth, M. R. (1983). *Labor supply*. Cambridge: Cambridge University Press.

Killingsworth, M. R., & Heckman, J. J. (1986). Female labor supply: A survey. *Handbook of Labor Economics, 1*(1), 103-204.

Kingdon, G., & Knight, J. (2007). Unemployment in South Africa, 1995–2003: Causes, problems and policies. *Journal of African Economies, 16*(5), 813-848.

Kingdon, G. G., & Knight, J. (2004a). Unemployment in South Africa: The nature of the beast. *World Development, 32*(3), 391-408.

Klasen, S., & Woolard, I. (1999). Levels, trends and consistency of employment and unemployment figures in South Africa. *Development Southern Africa, 16*(1), 3-35.

Knaul, F. M. (2001). Linking health, nutrition, and wages: The evolution of age at menarche and labor earnings among adult mexican women. *Inter-American Development Bank, 582*(4), 63-86.

Koenker, R., & Bassett Jr, G. (1978). Regression quantiles. *Econometrica: Journal of the Econometric Society, 46*(1), 33-50.

Kosters, M. H. (1966). Income and substitution effects in a family labor supply model. *The Rand Corporation.*

Kowalski, A. E. (2009). Censored quantile instrumental variable estimates of the price elasticity of expenditure on medical care. *NBER Working Paper 15085.*

Kreider, B. (1999). Latent work disability and reporting bias. *Journal of Human Resources,* , 734-769.

Kreider, B., & Pepper, J. (2003). *Disability and employment: Reevaluating the evidence in light of reporting errors,* Iowa State University, Department of Economics.

Lam, D. (1999). Generating extreme inequality: Schooling, earnings, and intergenerational transmission of human capital in South Africa and Brazil.

Lambrinos, J. (1981). Health: A source of bias in labor supply models. *The Review of Economics and Statistics, 63*(2), 206-212.

Larsson, D., Hemmingsson, T., Allebeck, P., & Lundberg, I. (2002). Self-rated health and mortality among young men: What is the relation and how may it be explained? *Scandinavian Journal of Public Health, 30*(4), 259-266.

LaRue, A., Bank, L., Jarvik, U., & Hetland, M. (1979). Health in old age: How do physicians' ratings and self-ratings compare? *Journal of Gerontology, 34*(5), 687-691.

Lee, L. F. (1982). Health and wage: A simultaneous equation model with multiple discrete indicators. *International Economic Review, 23*(1), 199-221.

Leibbrandt, M., Wegner, E., & Finn, A. (2011). The policies for reducing income inequality and poverty in South Africa. *SALDRU Working Paper Series, no. 64.*

Leibbrandt, M., & Woolard, I. (2001). The labour market and household income inequality in South Africa: Existing evidence and new panel data. *Journal of International Development, 13*(6), 671-689.

Leibbrandt, M., Finn, A., & Woolard, I. (2012). Describing and decomposing post-apartheid income inequality in South Africa. *Development Southern Africa, 29*(1), 19-34.

Leibenstein, H. (1957). *Economic backwardness and economic growth: Studies in the theory of economic development*. New York: John Wiley.

Levinsohn, J., McLaren, Z. M., Shisana, O., & Zuma, K. (2013). HIV status and labor market participation in South Africa. *The Review of Economics and Statistics, 95*(1), 98-108.

Levinson, D., Lakoma, M. D., Petukhova, M., Schoenbaum, M., Zaslavsky, A. M., Angermeyer, M., . . . Kessler, R. C. (2010). Associations of serious mental illness with earnings: Results from the WHO world mental health surveys. *The British Journal of Psychiatry : The Journal of Mental Science, 197*(2), 114-121. doi:10.1192/bjp.bp.109.073635 [doi].

Li, R., Chen, X., Yan, H., Deurenberg, P., Garby, L., & Hautvast, J. G. (1994). Functional consequences of iron supplementation in iron-deficient female cotton mill workers in beijing, china. *The American Journal of Clinical Nutrition, 59*(4), 908-913.

Loh, E. S. (1993). The economic-effects of physical appearance. *Social Science Quarterly, 74*(2), 420-438.

Long, C. D. (1958). The labor force under changing income and employment. *NBER Books.*

Lundberg, S. J. (1991). The enforcement of equal opportunity laws under imperfect information: Affirmative action and alternatives. *The Quarterly Journal of Economics, 106*(1), 309-326.

Lundberg, S., & Pollak, R. A. (1996). Bargaining and distribution in marriage. *The Journal of Economic Perspectives, 10*(4), 139-158.

Machado, J., & Mata, J. (2005). Counterfactual decomposition of changes in wage distributions using quantile regression. *Journal of Applied Econometrics, 20*(4), 445-465.

Madden, D. (2004). Labour market discrimination on the basis of health: An application to UK data. *Applied Economics, 36*(5), 421-442.

Manning, W. G., Newhouse, J. P., & Ware Jr, J. E. (1982). The status of health in demand estimation; or, beyond excellent, good, fair, poor. In V. R. Fuchs (Ed.), *Economic aspects of health* (pp. 143-181). Chicago, Illinois: University of Chicago Press.

Maranto, C. L., & Stenoien, A. F. (2000). Weight discrimination: A multidisciplinary analysis. *Employee Responsibilities and Rights Journal, 12*(1), 9-24.

Marcotte, D. E., & Wilcox-Gok, V. (2003). Estimating earning losses due to mental illness: A quantile regression approach. *Journal of Mental Health Policy and Economics, 6*(3), 123-134.

Marmot, M. (2002). The influence of income on health: Views of an epidemiologist. *Health Affairs, 21*(2), 31-46.

Mastrobuoni, G. (2009). Labor supply effects of the recent social security benefit cuts: Empirical estimates using cohort discontinuities. *Journal of Public Economics, 93*(11), 1224-1233.

McCord, A., & Bhorat, H. (2003). Employment and labour market trends. *Human Resources Development Review,* , 112-141.

McGarry, K. (2004). Health and retirement: Do changes in health affect retirement expectations?. *Journal of Human Resources, 39*(3), 25.

McIntyre, D., Gilson, L., Valentine, N., & Soderlund, N. (1998). *Equity of health sector revenue generation and allocation: A South African case study* Health Economics Unit, University of Cape Town.

Mckeown, T. (1976). *The modern rise of population* Edward Arnold London.

Melly, B. (2005). Decomposition of differences in distribution using quantile regression. *Labour Economics, 12*(4), 577-590.

Mete, C., & Schultz, P.T. (2002). Health and labour force participation of older people in Taiwan. *Economic Growth Center Discussion Paper, No. 846, Yale University.*

Mincer, J. (1974). Schooling, experience, and earnings. *Human Behavior & Social Institutions, 2.*

Mincer, J. (1962b). Labor force participation of married women. *Aspects of labor economics* (pp. 63-106) Princeton University Press.

Mitchell, J. M., & Burkhauser, R. V. (1990). Disentangling the effect of arthritis on earnings: A simultaneous estimate of wage rates and hours worked. *Applied Economics, 22*(10), 1291-1309.

Mitra, S. (2008). The recent decline in the employment of persons with disabilities in South Africa, 1998-2006. *South African Journal of Economics, 76*(3), 480-492.

Mitra, S., & Sambamoorthi, U. (2008). Disability and the rural labor market in india: Evidence for males in tamil nadu. *World Development, 36*(5), 934-952.

Mizunoya, S., & Mitra, S. (2013). Is there a disability gap in employment rates in developing countries? *World Development, 42*, 28-43.

Mortensen, D. (1986). Job search and labour market analysis. In O. Ashenfelter, & R. Layard (Eds.), *Handbook of labour economics* (1st ed., pp. 849-919). Amsterdam: Elsevier Science Publishers B.V.

Murray, C. J. (1994). Quantifying the burden of disease: The technical basis for disability-adjusted life years. *Bulletin of the World Health Organization, 72*(3), 429.

Murray, M. P. (2006). Avoiding invalid instruments and coping with weak instruments. *The Journal of Economic Perspectives, 20*(4), 111-132.

Murrugarra, E., & Valdivia, M. (1999). The returns to health for peruvian urban adults: Differentials across genders, the life-cycle and the wage distribution. *Inter-American Development Bank.*

Mushkin, S. J. (1962). Health as an investment. *The Journal of Political Economy, 70*(5), 129-157.

Mwabu, G., & Schultz, T. P. (1996). Education returns across quantiles of the wage function: Alternative explanations for returns to education by race in South Africa. *The American Economic Review, 86*(2), 335-339.

Mwabu, G., & Schultz, T. P. (2000). Wage premiums for education and location of South African workers, by gender and race. *Economic Development and Cultural Change, 48*(2), 307-334.

Nagi, S. Z. (1969). Congruency in medical and self-assessment of disability. *IMS, Industrial Medicine and Surgery, 38*(3), 27-36.

National Income Dynamics Study. (2013). *National income dynamics study: Wave 3 overview.* Pretoria: The Presidency, South Africa.

National Treasury. (2000). *Economic policy outlook.* Pretoria: National Treasury.

National Treasury. (2008). *Economic policy and outlook.* Pretoria: National Treasury.

National Treasury. (2014). *Economic outlook.* Pretoria: National Treasury.

Nattrass, N., & Walker, R. (2005). Unemployment and reservation wages in working-class Cape Town. *South African Journal of Economics, 73*(3), 498-509.

Neuman, S., & Oaxaca, R. L. (2004). Wage decompositions with selectivity-corrected wage equations: A methodological note. *Journal of Economic Inequality, 2*(1), 3-10.

Neuman, S., & Oaxaca, R. L. (2003). Estimating labor market discrimination with selectivity-corrected wage equations: Methodological considerations and an illustration from israel. *Discussion Paper no. 2-2003.*

Neumark, D. (1988). Employers' discriminatory behavior and the estimation of wage discrimination. *Journal of Human Resources, 23*(3), 279-295.

Novignon, J., Novignon, J., & Arthur, E. (2015). Health status and labour force participation in Sub-Saharan Africa: a dynamic panel data analysis. *African Development Review, 27(1), 14-26.*

Ntuli, M. (2009). *Exploring the status of African women in the South African labour market, 1995-2004.* (Unpublished Ph.D.). University of Cape Town, Cape Town.

Ntuli, M., & Wittenberg, M. (2013). Determinants of black women's labour force participation in post-apartheid South Africa. *Journal of African Economies, 0*(0), 1-28.

Nybo, H., Petersen, H. C., Gaist, D., Jeune, B., Andersen, K., McGue, M., . . . Christensen, K. (2003). Predictors of mortality in 2,249 Nonagenarians—The danish 1905-Cohort survey. *Journal of the American Geriatrics Society, 51*(10), 1365-1373.

O'Donnell, O. (1998). The effect of disability on employment allowing for work incapacity. *University of Kent at Canterbury Department of Economics Working Paper Series, 46*(98-13).

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review, 14*(3), 693-709.

Oettinger, G. S. (1996). Statistical discrimination and the early career evolution of the black-white wage gap. *Journal of Labor Economics, 14*(1), 52-78.

Ogloblin, C. G. (1999). The gender earnings differential in the russian transition economy. *Industrial and Labor Relations Review, 52*(4), 602-627.

O'Hara, B. (2004). Twice penalized employment discrimination against women with disabilities. *Journal of Disability Policy Studies, 15*(1), 27-34.

Oppenheimer, V. K. (1997). Women's employment and the gain to marriage: The specialization and trading model. *Annual Review of Sociology, 23*, 431-453.

Palmer, M. G., Thuy, N. T. M., Quyen, Q. T. N., Duy, D. S., Van Huynh, H., & Berry, H. L. (2012). Disability measures as an indicator of poverty: A case study from viet nam. *Journal of International Development, 24*(S1), S53-S68.

Parker, S. W., & Knaul, F. (1999). *Elderly health and salaries in the mexican labor market* Inter-American Development Bank, Office of the Chief Economist.

Parsons, D. O. (1980). The decline in male labor force participation. *The Journal of Political Economy, 88*(1), 117-134.

Parsons, D. O. (1982). The male labour force participation decision: Health, reported health, and economic incentives. *Economica, 49*(193), 81-91.

Patterson, I., & Pegg, S. (1999). Nothing to do. the relationship between leisure, boredom and alcohol and drug addiction: Is there a link to youth suicide in rural Australia? *Youth Studies Australia, 18*(2), 24-29.

Pelkowski, J. M., & Berger, M. C. (2004). The impact of health on employment, wages, and hours worked over the life cycle. *The Quarterly Review of Economics and Finance, 44*(1), 102-121.

Pencavel, J. (1986). Labour supply of men: A survey. In O. Ashenfelter, & R. Layard (Eds.), *Handbook of labour economics* (1st ed., pp. 3-102). Amsterdam: Elsevier Science Publishers B.V.

Phelps, E. S. (1972). The statistical theory of racism and sexism. *The American Economic Review, 62*(4), 659-661.

Posel, D., & Casale, D. (2003). What has been happening to internal labour migration in South Africa, 1993–1999? *South African Journal of Economics, 71*(3), 455-479.

Powell, J. L. (1986a). Censored regression quantiles. *Journal of Econometrics, 32*(1), 143-155.

Prince, M., Patel, V., Saxena, S., Maj, M., Maselko, J., Phillips, M. R., & Rahman, A. (2007). No health without mental health. *The Lancet, 370*(9590), 859-877.

Pritchett, L., & Summers, L. H. (1996). Wealthier is healthier. *Journal of Human Resources,* , 841-868.

Ranchhod, V. (2009). *Labour market: Analysis of the NIDS wave 1 dataset.* (No. 12).

Register, C. A., & Williams, D. R. (1990). Wage effects of obesity among young workers. *Social Science Quarterly, 71*(1), 130-141.

Reimers, C. W. (1983). Labor market discrimination against hispanic and black men. *The Review of Economics and Statistics, 65*(4), 570-579.

Ribero, R., & Nuñez, J. (2001). Productivity of household investment in health: The case of Colombia. *Investment in Health: Social and Economic Returns, 582*(4), 35-62.

Rivera, B., & Currais, L. (Eds.). (2005). *Individual returns to health in Brazil: A quantile regression analysis* The MIT Press.

Robbins, L. (1930). On the elasticity of demand for income in terms of effort. *Economica,* (29), 123-129.

Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *The Journal of Political Economy, 82*(1), 34-55.

Rosenzweig, M. R., & Wolpin, K. I. (1986). Evaluating the effects of optimally distributed public programs: Child health and family planning interventions. *The American Economic Review, 76*(3), 470-482.

Rospabé, S. (2002). How did labour market racial discrimination evolve after the end of apartheid? *South African Journal of Economics, 70*(1), 185-217.

Rospabe, S., & Selod, H. (2006). Does city structure cause unemployment? the case of cape town. In H. Bhorat, & R. Kanbur (Eds.), *Poverty and policy in post-apartheid South Africa* (pp. 262-287) HSRC (Human Sciences Research Council) Press Cape Town.

Ross, M. H., & Murray, J. (2004). Occupational respiratory disease in mining. *Occupational Medicine, 54*(5), 304-310. doi:10.1093/occmed/kqh073 [doi].

SALDRU. (1994). South Africans rich and poor: Baseline household statistics. *School of Economics, University of Cape Town. Cape Town.*

Saracci, R. (1997). The world health organisation needs to reconsider its definition of health. *Bmj, 314*(7091), 1409.

Savedoff, W. D., & Schultz, T. P. (2000). *Earnings and the elusive dividends of health* Inter-American Development Bank.

Schultz, T. P. (2002). Wage gains associated with height as a form of health human capital. *American Economic Review, 92*(2), 349-353.

Schultz, T. P. (2004). Evidence of returns to schooling in Africa from household surveys: Monitoring and restructuring the market for education. *Journal of African Economies, 13*(suppl 2), ii95-ii148.

Schultz, T. P. (1988). Heterogeneous preferences and migration: Self selection, regional prices and programs, and the behavior of migrants in Colombia. *Research in Population Economics, 6*, 163-181.

Schultz, T. P. (2005). Productive benefits of health: Evidence from low-income countries. *Health and Economic Growth: Findings and Policy Implications.MIT Press, Cambridge MA,* , 257-286.

Schultz, T. P., & Mwabu, G. (1998b). Labor unions and the distribution of wages and employment in South Africa. *Indus.& Lab.Rel.Rev., 51*(4), 680-703.

Schultz, T. W. (1979). *Prize lecture.lecture to the memory of Alfred Nobel. (*Unpublished manuscript).

Seekings, J. (2008). The continuing salience of race: Discrimination and diversity in South Africa. *Journal of Contemporary African Studies, 26*(1), 1-25.

Seekings, J. (2014). Taking disadvantage seriously: The ‘underclass’in post-apartheid South Africa. *Africa, 84*(01), 135-141.

Smith, J. P. (2009). The impact of childhood health on adult labor market outcomes. *The Review of Economics and Statistics, 91*(3), 478-489.

Smith, J. P., & Smith, G. C. (2010). Long-term economic costs of psychological problems during childhood. *Social Science & Medicine, 71*(1), 110-115.

Solow, R. M. (1956). A contribution to the theory of economic growth. *The Quarterly Journal of Economics, 70*(1), 65-94.

South African Reserve Bank (undated). Online download facility, accessed on 30/05/2015, <https://www.resbank.co.za/Research/Statistics/Pages/OnlineDownloadFacility.aspx >

Spence, M., & Lewis, M. A. (2009). *Health and growth* World Bank Publications.

Sprague, A. (1994). Work experience, earnings and participation: Evidence from the women and employment survey. *Applied Economics, 26*(7), 659-667.

Staiger, D., & Stock, J. H. (1997). Instrumental variables regression with weak instruments. *Econometrica, 65*(3), 557-586.

Statistics South Africa. (2004). *Census 2001: Concepts and definitions.* (No. 03-02-26 (2001)). Pretoria: Statistics South Africa.

Statistics South Africa (2015). CPI history, accessed on 30/05/2015, <http://www.statssa.gov.za/publications/P0141/CPIHistory.pdf>

Statistics South Africa. (2012a). *Census 2011 statistical release.* (No. P0301.4). Pretoria: Statistics South Africa.

Statistics South Africa. (2012c). *Income and expenditure of households 2010/2011.* (No. P0100). Pretoria: Statistics South Africa.

Stern, S. (1989). Measuring the effect of disability on labor force participation. *Journal of Human Resources,* , 361-395.

Stigler, G. (1961). The economics of information. *The Journal of Political Economy, 69*, 213-225.

Stigler, G. (1962). Information in the labour market. *Journal of Political Economy, 70*, 94-105.

Strauss, J. (1986). Does better nutrition raise farm productivity? *The Journal of Political Economy, 94*(2), 297-320.

Strauss, J., & Thomas, D. (1998). Health, nutrition, and economic development. *Journal of Economic Literature, 36*(2), 766-817.

Suhrcke, M., McKee, M., Mortensen, J., Arce, R. S., & Tsolova, S. (2005). *The contribution of health to the economy in the European Union* Office for Official Publications of the European Communities.

Thomas, D., & Frankenberg, E. (2002). Health, nutrition and prosperity: A microeconomic perspective. *Bulletin of the World Health Organization, 80*(2), 106-113.

Thomas, D., Frankenberg, E., Friedman, J., Habicht, J., Hakimi, M., Jaswadi, N. J., . . . Smith, J. P. (2003). Iron deficiency and the well-being of older adults: Early results from a randomized nutrition intervention. *Population Association of America Annual Meetings, Minneapolis.*

Trani, J., & Loeb, M. (2012). Poverty and disability: A vicious circle? evidence from Afghanistan and Zambia. *Journal of International Development, 24*(S1), S19-S52.

UNDP. (2011). *Human development report - sustainability and equity: A better future for all.* New York: Palgrave, Macmillan.

UNDP. (2014). *Human development report - sustaining human progress: Reducing vulnerabilities and building resilience.* New York: United Nations Development Programme.

UNICEF (2010). Di Data Wizard, accessed on 30/05/2015 <http://www.devinfo.info/southafrica/DIWizard/DIWizardPreviews.aspx>

van Solinge, H., & Henkens, K. (2010). Living longer, working longer? the impact of subjective life expectancy on retirement intentions and behaviour. *European Journal of Public Health, 20*(1), 47-51. doi:10.1093/eurpub/ckp118 [doi].

Weil, D. N. (2007). Accounting for the effect of health on economic growth. *The Quarterly Journal of Economics, 122*(3), 1265-1306.

WHO. (1946). Preamble to the constitution of the World Health Organization as adopted by the International Health Conference, New York, 19-22 June, 1946; signed on 22 July 1946 by the representatives of 61 states (Official Records of the World Health Organization, no. 2, p. 100) and entered into force on 7 April 1948.

WHO. (1980). *The international classification of impairments, disabilities and handicaps.* Geneva: WHO.

WHO. (2001). Report of the commission on macroeconomics and health. *WHO.*

WHO. (2009). *Mortality and burden of disease estimates for WHO member states in 2004.*

Wilkinson, R. B. (1995). Changes in psychological health and the marital relationship through childbearing: Transition or process as stressor? *Australian Journal of Psychology, 47*(2), 86-92.

Winter, C. (1999). Women workers in South Africa: Participation, pay and prejudice in the formal labour market. *African Region Country Department. World Bank.*

Wittenberg, M. (1999). A spatial analysis of unemployment. *Symposium on Spatial Guidelines for Infrastructure, Investment and Development. Office of the Deputy President. Pretoria.*

Wittenberg, M. (2011). The weight of success: The body mass index and economic well-being in South Africa. *SALDRU Working Papers no.65.*

Wittenberg, M. (2014). *Analysis of employment, real wage, and productivity trends in South Africa since 1994.* (Conditions of Work and Employment Series No. 45). Geneva: International Labour Organization.

Wolfe, B. L., & Hill, S. C. (1995). The effect of health on the work effort of single mothers. *Journal of Human Resources, 30*(1), 42-62.

Wong, D. D., Wong, R. P., & Caplan, G. A. (2007). Self-rated health in the unwell elderly presenting to the emergency department. *Emergency Medicine Australasia, 19*(3), 196-202.

Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data.* Cambridge, MA: MIT Press.

Woytinsky, W. S. (1940). *Additional workers and the volume of unemployment in the depression.* New York: Social Science Research Council.

Young, A. (2005). The gift of the dying: The tragedy of AIDS and the wealth of future African generations. *Quarterly Journal of Economics, 120*(2), 423-466.

Zhang, W., O'Brien, N., Forrest, J. I., Salters, K. A., Patterson, T. L., Montaner, J. S., . . . Lima, V. D. (2012). Validating a shortened depression scale (10 item CES-D) among HIV-positive people in British Columbia, Canada. *PloS One, 7*(7), 1-5.

Zweifel, P., Breyer, F., & Kifmann, M. (2009). *Health Economics,* Springer.