LUND UNIVERSITY
School of Economics and Management

Health, Income and Local Comparisons

A Study of the Relative Income Hypothesis in South Africa

by:

Niklas Wallin Bernhardsson & David Wästlund

August 2015

Master's Programme in Economics

Supervisor:Ulf Gerdtham

Abstract

This study investigates the relationship between income and health of individuals in South Africa. In particular it tests whether the relative, rather than just the absolute, income of the individual plays a significant part in determining health outcomes. The absolute income hypothesis (AIH) is tested and, in line with previous related research, found significant, so that individuals with higher income are associated with better health. The Relative Income Hypothesis (RIH) is then tested through the added inclusion of a term for the individual's relative income in his or her community.

The study uses Self-Assessed Health (SAH) from a nationally representative sample to determine individual health. It uses two different types of measures for relative income, one objective based upon reported income and one subjective based upon the individual's perception of how the own income relates to the income of others. The role of knowledge and groups for comparison available to the individual is thoroughly discussed. The Relative Income Hypothesis (RIH) is also separately tested for different income groups as well as across time, since both past and predicted future income may be influential in determining current health.

The study finds that whether relative income helps explaining individual health strongly depends on the measure used for relative income. In particular whether relative income is defined in objective or subjective terms is found to be crucial. The results based on objective income finds no support for the Relative Income Hypothesis whereas support is found when subjective income is used. This suggest the importance of psychological factors in explaining the relationship between income and health.

Keywords: Relative Income Hypothesis, Self-Assessed Health, Absolute Income Hypothesis, South African Health

Table of Contents

1 - Introduction 1
1.1 Purpose of study 2
1.1.2 Sub-questions 2
1.1.3 Limitations 2
1.1.4 Contribution to the academic field 3
1.1.5 Disposition 3
2 - Background 4
2.1 South Africa 5
3 - Theory 7
3.1 The absolute income hypothesis 7
3.2 The relative income hypothesis 8
3.2.1 Through which mechanisms can relative income affect health? 9
3.2.2 Relative income hypothesis for different income groups 10
3.3 Spill-overs on a local level 10
3.4 Health, income and time 11
4 - Method 13
4.1 General Intuition 13
4.2 The two measures of relative income 14
4.3 Model specification for the objective model 15
4.4 Creating reference groups 16
4.5 Model specification for the subjective model 17
4.6 Estimation method 18
4.7 Further method considerations 19
4.7.1 Does the RIH differ across income groups? 19
4.7.2 Intertemporal correlation between relative income and health 20
4.7.3 Alternative measures of absolute income 20
5 - Data 21
5.1 About the data 21
5.2 Health 22
5.3 Income 23
5.4 Other variables 25
5.5 Descriptive data 27
6 - Results 30
6.1 Main Results 30
6.2 Relative income hypothesis for different income groups 32
6.3 Intertemporal correlation 33
6.4 Robustness checks 35

7 - Analysis & Discussion 37
8 - Conclusion 42
9 - References 43
Appendix A – List of Variables 46
Appendix B- Additional figures and graphs 48

1 - Introduction

Although the correlation between income and health is one of the most established within health economics much about this connection is yet to be understood. The Absolute Income Hypothesis (AIH) has emerged as the main theory, relating the health status of the individual to the individual's income. The theory has found strong empirical support from different data materials and methods, both between and within countries (see e.g. Wagstaff & van Doorslaer, 2000; Karlsson et al, 2010).

Despite the success of the AIH, it is debated whether it can really explain all of the correlation between income and health. Several competing theories regarding this relationship have been established in the academic literature. Among them, the Relative Income Hypothesis (RIH) has been prominent. Unlike the AIH, the RIH factor in the income, not just of the individual, but for the individual's entire community. The underlying idea is that what matters is not only the absolute income of the individual but also how that income relates to the income of others. An implication of the theory is that an individual would do better, health-wise, being relatively rich in a poorer community than relatively poor in a wealthier one, even should the absolute income be the same.

Where an ample amount of academic literature discusses the effects of income distribution and inequality on health, the question regarding if changes to an individual's relative income has health impacts has received surprisingly little attention. Empirical research has mostly been conducted in the developed world and in particular in the US. Findings have been mixed: some studies have supported the RIH, others have failed to do so and some have even outright contradicted it (Wagstaff & van Doorslaer, 2000). However, studies have shown that the picture may be different for developing countries and newly industrialized countries where the correlation between income and health is generally stronger (Fuchs, 2004; Karlsson et al, 2010).

As the empirical research over the Relative Income Hypothesis using individual-level data on developing countries and newly industrialized countries remains scarce, this study hopes to contribute and drive this research forward. To further test the RIH in a newly industrialized country, South Africa is considered. With a growing body of available data over the

population and the presence of substantial income inequality, South Africa appears to be well suited for this purpose.

This paper will investigate the relevance of the relative income hypothesis on the health outcomes of the South African population. This is done by first establishing the health effects from absolute income and then test whether the relative income hypothesis might offer additional explanatory value for the health status.

The relative income hypothesis is a framework to understand the impact of income and social-inequalities on individual health. Knowing whether relative income matters may have important implications for the targeting of both health and income distribution policies. On top of this, if the RIH holds true, health may be dependent upon subjective income comparisons. This would make it important for society to understand who individuals compare their own income to and how more suitable reference groups could be encouraged. By finding if relative income matter, this study hope to contribute to a general understanding and improving of income-related health.

1.1 Purpose of study

This study will investigate if the Relative Income Hypothesis (RIH) is relevant for explaining health status in South Africa.

1.1.2 Sub-questions

- Is the absolute income of the individual relevant in explaining health outcomes?
- Is there a difference between how objective and subjective relative income affects health?
- Does the RIH differ across income groups?
- Is there an intertemporal relationship between relative income and health?

1.1.3 Limitations

This study does not attempt to prove or investigate the direction of causality for relative income and health. It solely considers whether the two are significantly correlated. Neither does it attempt to find all the relevant determinants of health, nor estimate their exact size.

The individual's subjective health status is the only health measure under consideration, specific health issues are not explored. This study applies the definition of the RIH according to Wagstaff & van Doorslaer (2000), consequently this study will not investigate other income effects on health as for example income inequality and relative deprivation.

Despite their names, both the Absolute Income Hypothesis (AIH) and the Relative Income Hypothesis (RIH) are as applicable to wealth as they are to income. This study measures them both, although income is used in some contexts and wealth in others depending on the data available. However, the conclusions drawn regarding health effects are the same, regardless of whether income or wealth is measured.

Within this paper health outcomes for children are not included. Children are defined as people in the household below 15 years of age.

1.1.4 Contribution to the academic field

This paper adds to the literature on the Relative Income Hypothesis (RIH) in newly industrialized countries by studying whether the RIH holds true for South Africa. It uses individual-and household-data on both income and health from a nation-wide survey. The study also uses data on subjective relative income to test the RIH. To the best of knowledge, it is unique in combining both subjective and objective measures on relative income in the same study on health. This makes it possible to not only test the relative income hypothesis as a whole but also to distinguish between individual effects within the hypothesis. Furthermore, as the relevance of the RIH might differ across income groups, this study will test if such differences also exist in newly industrialized countries. A final contribution of the study is that it also briefly examines intertemporal correlation between relative income and health, thus shining more light on through which mechanisms the RIH may work.

1.1.5 Disposition

This paper is structured as follows: Chapter 2 will give a brief background to previous research on the relative income hypothesis as well as health and health care in South Africa. Chapter 3 will outline the theories behind the absolute and relative income hypothesises. Chapter 4 will describe the methods used to answer the research question while Chapter 5 will present information about the data used. The results will then be presented in Chapter 6 and analysed in Chapter 7 before Chapter 8 summarizes the conclusions of the study.

2 - Background

The Relative Income Hypothesis (RIH) emphasize that it is not only the direct effect of income that impact the health status of the individual, but also the size of the income in relation to that of others. These effects are theorized in the RIH to work through multiple mechanisms, for example through psychological factors and the availability of health goods. The RIH states that with a constant income of the individual, changes in the reference group or the surroundings will influence the individuals' health. Phrased differently, the theory states that the same absolute income will yield better health for an individual living in a poorer community than for an identical person living in a wealthier one (Wagstaff & van Doorslaer, 2000).

Several studies have found evidence that the RIH is relevant in explaining health, well-being and happiness (e.g. McBride, 2001; Clark et al, 2008). However, many other studies on the topic has found no evidence of the RIH, or even outright contradicted it (Wagstaff & van Doorslaer, 2000). On top of this, a large share of the research conducted have been at the population-level where comparison studies are performed over different countries, where even though some results support the RIH, a consensus among researchers have not been reached, as the methods have been criticized (Ellison, 2002). To complicate matters further, it has been found that the relevance of relative income may differ between income groups and across countries (Karlsson et al, 2010).

When the RIH first emerged the definitions varied between studies. A common idea was that income inequality of a society was in itself negatively correlated with individual health. This idea has received much attention and many studies have tested it with various results (see e.g. Wagstaff & van Doorslaer, 2000; Mellor & Milyo, 2002; Karlsson et al, 2010). This study however follows the more strict definition of the RIH applied by Wagstaff & van Doorslaer (2000). Following this definition, all direct effects to health from income inequality is referred to as the Income Inequality Hypothesis (IIH) and these effects will not be tested in this study. Compared to the IIH, the stricter definition of the RIH used in this study has so far only received modest attention in the health economic literature. Clearly, more empirical evidence based upon individual-level data is needed in order to fully understand how relative income affects individual health.

2.1 South Africa

The context in which the Relative Income Hypothesis (RIH) is tested has through earlier studies been proven to have large impact on its relevance (see. e.g. Karlsson et al, 2010). Consequently an introduction of South Africa is necessary before proceeding to test the RIH. A brief introduction to some general key features regarding the population's health and the South African health care system will also be provided.

South Africa is a multiethnic newly industrialized country with a long history of apartheid, social stratification and income inequality. It is a nation with multiple languages, religions and cultures, which are wide-spread across its 9 different provinces. These provinces are in turn divided into 52 different districts (Brand South Africa, 2015). The colonial history with the following apartheid has created inequality in income distribution and segregation between ethnic groups and cultures. Even though the first democratic election was held in 1994, the South African society still suffers from segregation and huge income differences across the population (CIA World Factbook [a], 2015).

The country has one of the world's most unequal income distributions (CIA World Factbook [b], 2015). The standard way of measuring income distribution is the GINI-index which ranges from 0 to 1 where a higher number indicates more inequality. In 2011 South Africa had a Gini coefficient of 0,65 as compared to e.g. the United states of America which in 2010 had a coefficient of 0,411 (World Bank, 2015).

The quality of health care naturally has an impact on population health. However, the financing and distribution of health care is also important and crucial to assess the importance of relative income. The South African health care system is managed in three different levels, with the South African Health department as the top institution that mostly focus on the public sector. The management and supervision is conducted at the province-level and has an overview over the district-based and local hospitals, which is providing the actual health care and makes decision regarding its day-to-day business (Coovadia et al, 2009).

The South African health care system consists of both a large public sector and a private health care system where the latter is mostly utilized by middle- and high-income earners. The public sector is contributing with about 80 % of all the health care provided to the South

African population. The quality of the public health care is poor as a result of the lack of important resources both in funding and in competence in forms of doctors and technology. There is a scarcity of educated doctors in South Africa, which contributes to the difficulty of providing quality health care to the citizens, and out of these educated doctors 73 % work in the private sector. Consequently there is an additional divergence in doctor-citizens ratio between public and private sector which is reflected in the quality differences of health care (Brand South Africa, 2015). Therefor the economic status and geographical location of the household are important determines of the health care provided.

The average health status in South Africa is quite poor which is reflected by the low life expectancy of 59 years in 2012. One of the biggest health problems in South Africa is the widespread HIV/AIDs which by WHO´s estimation in 2012 was responsible for 33,2 % of all deaths. The HIV/AIDS presence is wide-spread across South Africa and its prevalence is estimated to be 10.6 % of South Africans population. This can have both mortal effects and effects on the average health status of the population, partly depending to the treatment of the infected. This emphasizes the importance of a functioning health care system offered to all citizens of South Africa (WHO, 2015).

What is important for the reader to take from this section is that South Africa differs considerably from the developed world regarding some key aspects. Income distribution is highly skewed towards the wealthiest individuals and the health care system is constructed so that health care quality increases rapidly with the ability to pay. Due to this, findings may differ substantially from most previous studies on the Relative Income Hypothesis (RIH) and comparisons needs to be done with caution.

3 - Theory

There are many theories regarding in which ways health can be affected by income. Two important ones are the Absolute Income Hypothesis (AIH) which relates the income and health of the individual to each other, and the Relative Income Hypothesis (RIH) which on top of this also expect health to be correlated with the income of others. This section will present both of these theories. After this it will explore spill-over effects on health from income of the local community. It will conclude with a discussion of intertemporal correlations between income and health.

3.1 The absolute income hypothesis

A clear correlation has been found between wealth and health of the individual. This may be explained by three things. First of all, there may be direct health benefits from having more money. This seems realistic if a higher budget allows more spending on health goods, something predicted in the basic Grossman model (Grossman, 1972). Second, being healthier may lead to higher wealth. This makes sense as a healthy person has better prospects of working and thereby earn money. Third, there may be yet another factor that causes both better health and higher wealth. In practice, a combination of all three is likely to be true in explaining the high positive correlation between income and health (Pritchett & Summers, 1997).

The positive correlation between health and wealth/income is found for virtually all countries (e.g. Karlsson et al, 2010) and studies (Wagstaff & van Doorslaer, 2000). The correlation also holds true for many different health measures, such as Self-Assessed Health (SAH) (e.g. Subramanian et al, 2001), mortality (e.g. Gerdtham & Johannesson, 2004) and life-expectancy (Pritchett & Summers, 1997). The idea that income is positively correlated to better health is often referred to as the Absolute Income Hypothesis (AIH). The relationship is usually thought of as being concave, so that the marginal impact on health diminishes with a rising income (Wagstaff & van Doorslaer, 2000). The correlation is thought to be stronger for countries with below-average income (Fuchs, 2004).

The AIH has also frequently been used to explain the ties between income inequality and health. Due to the seemingly negative relationship between income inequality and health a number of alternative theories have been proposed, such as the income-inequality hypothesis. It has however been shown that the AIH itself is able to explain much of this

negative correlation (Gravelle, 1998). Despite the theory's simplicity, it has remained the predominant theory for explaining the relationship between both income and health and income inequality and health (Wagstaff & van Doorslaer, 2000).

The relationship between income levels and health outcomes has been well-established in both empirical findings and in theory. There are however challenges in detangling any causal effects due to the above mentioned problem with reversed causality (Case, 2004; Erdil & Yetkiner, 2004, Fuchs, 2004). Since this study does not attempts to establish causality it is enough to know that a significant correlation between income and health is expected.

3.2 The relative income hypothesis

The idea behind the Relative Income Hypothesis (RIH) is that the income in relation to the income of others impacts the individual's health beyond just the income in absolute terms. This is an important distinction as in the Absolute Income Hypothesis (AIH) income levels for the rest of the population has no correlation with the health of the individual. The RIH hence expects a positive correlation between relative income and the health of the individual, so that an individual earning less than his or her peers can expect negative health effects and an individual earning more can expect positive health outcomes. Phrased differently, the RIH predicts a negative correlation between the income of peers and the health of the individual, as higher income for everyone else would make the individual relatively poorer – something associated with worse health (Wagstaff & van Doorslaer, 2000; Kawachi et al, 2002).

In some stricter version of the RIH, relative income is the only way in which income is expected to matter. Taken to the extreme, this would mean that if income was doubled for every person in society there would be no impact on overall health levels. In this study a less strict version of the RIH is however used so that changes to absolute income could affect the health of the individual. Note that this definition of the RIH still differs from the AIH, since the AIH predicts no correlation between relative income and individual health.

One crucial point of the RIH is that it is income compared to one's peers that matters, not income in relation to the general community. Because of this, even individuals with high absolute income may suffer from being relatively poor if their peers are better off than they are. (Eibner & Evans, 2005). As a result of this, the effect of income may affect individuals differently depending on whom they compare their own income to. Consequently knowing

whom individuals compare their own income to is crucial when testing for the RIH. The group to which income is compared can be called the individual's reference group.

The academic literature is divided regarding which groups constitute appropriate reference groups. Suggestions regarding the basis for reference groups include for example ethnicity/race and town type (Kawachi et al, 2002). However, basing reference groups upon geography is a common practice, as it seems reasonable that the perception of one's own income is influenced by the income of other people in the local community. Although there are no clear guidelines regarding the appropriate size of the geographical division, Jones, Duncan & Twigg (2004) find evidence that in the UK, districts make better reference group than regions. They further propose that using even smaller geographical units would be preferable should the data allow it.

The RIH expects a positive relationship between relative income and individual health, both when the individual is earning more or less money than his or her peers. However, both the reasons for this and the magnitude of the impact may differ slightly depending on if the individual is relatively rich or poor (Subramanian et al (2001) and Blanco-Perez (2012)). Health impacts are thought to be particularly severe for those individuals feeling poorer than their peers (Eibner & Evans, 2005).

3.2.1 Through which mechanisms can relative income affect health?

The academic literature on the topic displays two main mechanisms through which the relative income of the individual can affect the individual's health status. These two mechanisms are briefly outlined below.

The relative size of income compared to society as a whole may have direct effects on the health of the individual. Being relatively poor is expected to raise psychosocial stress which in turn can lead to worse health (Grönqvist et al, 2012). Relative income is closely related to socioeconomic status and financial stress, both of which can affect an individual's health either directly, e.g. through increased blood pressure, or indirectly, e.g. through changed behaviour such as increased smoking and drinking (Eibner & Evans, 2005). Comparing the own income with the income of one's peers may hence have health consequences through affecting psychological well-being. In this line of reasoning, the feeling of being rich or poor may matter more than actually being rich or poor (Miller & Paxson, 2006). It has also been

suggested that being relatively poor may result in anger and resentfulness towards society, something that is also associated with worse health status (Wilkinson, 1999).

Relative income may also affect health through the availability of different health goods. Holding the individual's income constant while changing the income of everyone else may impact upon individual health through price changes. Assume that the general price level of health care, nutritious food and general health goods would change as a result from everyone else's income having changed. This might greatly affect the individual's ability to manage the own health, according to the basic Grossman model (Grossman, 1972). These relative price effects imply a negative correlation between community income and the health of the individual (Miller & Paxson, 2006). Naturally this line of reasoning can also be applied without price changes. If e.g. the cost of medical care remained unchanged in a community whilst the average income increased, more people would be able to afford an unchanged supply of medical care. This increased demand would risk adversely affecting individuals who could more easily have received medical care before the wage rises.

3.2.2 Relative income hypothesis for different income groups

There is evidence from previous research that the Relative Income Hypothesis (RIH) is more present amongst some income groups than others. The heterogeneity of relative income across income groups is established in Subramanian et al (2001) study over how the RIH impacts on objective health measures at the state-level in the US. This is also confirmed by McBride (2001), which further finds that absolute income increases have larger effect on low-income groups than for high-income groups, suggesting a marginally decreasing impact on health from income. However, for relative income the effect of an increase in reference group income is found larger among high-income groups. If correct, this means that the RIH matters more for high-income groups while it adds little predictive value on top of the AIH among low-income groups.

3.3 Spill-overs on a local level

The Relative Income Hypothesis (RIH) predicts health to be negatively correlated with the income of others. However, local spill-overs may have the opposite effect. This is the case if re-distributional forces positively affect the health of the individual without affecting his or her income. For instance, an individual can benefit from higher income of other people if this results in higher tax income to finance public health care (Miller & Paxson, 2006). It could be

noted that the opposite effect also has been suggested, that increased community income could result in less public health care if it would make people more reluctant towards financing health care through taxes (ibid.). Naturally, the validity of this argument differs between different communities and cultures.

Local spill-overs may arise even without re-distributional forces. Other neighbourhood factors could impact upon the health of the individual. Wealthier neighbourhoods often display a safer environment, less pollution and a more health-aware behaviour in general (Coovadia et al, 2009). The effect of these factors might lead to that the income of others could be positively correlated to individual health.

Even if local spill-overs may lead to a positive correlation between individual health and the income of others – the opposite of what the RIH predicts – it does not invalidate the RIH. Relative income may still be affecting the health of the individual through psychological factors and price effects, even though the net effect is mitigated (or potentially even eroded or reversed) by local spill-overs. This however does make it more difficult to capture the effect of the RIH through its different mechanisms.

As should be clear from this discussion, the net impact of relative income on individual health is hard to predict. There is no theoretical consensus as to whether the net effect of these factors would lead to a positive or negative correlation with individual health. Although the relative income hypothesis predicts a positive correlation between relative income and individual health, the effect could theoretically be the opposite if the spill-over effects would outweigh the effects proposed by the RIH.

3.4 Health, income and time

Both the Absolute Income Hypothesis (AIH) and the Relative Income Hypothesis (RIH) suggests a correlation between income and health. However, they do not state that health solely exhibits a correlation with current income. Both past and predicted future income may also be significant in explaining current health, perhaps even more so than current income. There are many potential explanation for this. First of all, income may have a lagged effect on health, so that current health is a result of (accumulated) past income rather than current one. This has been found true by e.g. Benzeval & Judge (2001). Psychological factors

may also play its part, so that future health and income are affected by subjective perceptions of current life situation.

The discussion on how the correlation changes over time is especially interesting to studies trying to determine the direction of causality between health and income. This in turn is important for some policy decisions. If income determines health, income distribution policies are a way of changing societal health. Income could affect health, e.g. through the possibility to invest in the own health capital (Grossman, 1972). If by contrast health determines income, societal health policies would also impact on the economy. Health could affect income since it determines how much and how long people can work in their lives, as well as their productivity. A two-way causality has been found, so that income affects health but that health simultaneously affects income, something that has repeatedly caused trouble for studies trying to determine the exact size of either of these causalities (Erdil & Yetkiner, 2004). Naturally health and income may also be affected by unobserved third factors (Fuchs, 2004).

Causality patterns may of course differ, depending for instance on the individual's country, age and income group. Erdil & Yetkiner (2004) have found causality from income to health to be stronger for low income countries than in the developed world.

4 - Method

The purpose of this study is to test if the Relative Income Hypothesis (RIH) is relevant for explaining health status in South Africa. As stated above, the RIH predicts individual health to be positively correlated with relative income. To test if this holds true, the study regresses Self-Assessed Health (SAH) upon a term for relative income, as well as other determinants on health. This chapter explains the ideas behind such regressions, how they are performed and how the measures of health and relative income are created.

This chapter will start by explaining the general intuition behind testing the RIH and why finding a good measure of absolute income is crucial. The study uses two different types of measures for relative income and both the intuition behind this and the respective model specifications are explained in the following sections. After this the econometric estimation method used is presented. Finally the chapter concludes by presenting some further tests that are relevant to establish the connections between relative income and health.

4.1 General Intuition

The Relative Income Hypothesis (RIH) predicts a correlation between relative income and health. To test if this holds true, the study regresses Self-Assessed Health (SAH) upon a term for relative income, as well as other determinants on health.

Naturally a good measure for relative income is crucial in testing the RIH. However, finding a measure of relative income is not without its problems. This is because having a relatively high income is correlated with also having a high income in absolute terms. If the only thing tested is the correlation between SAH and relative income, the relevance of the RIH would be exaggerated since such a correlation would partly consist of the correlation between health and absolute income. Such a test would hence be unable to test if the RIH adds further explanatory value on top of the AIH. To be able to detangle health effects from relative and absolute income, terms for both need to be entered into the same model.

In its most basic form, a test for the Absolute Income Hypothesis (AIH) can be based on equation 1. By regressing Self-Assessed Health (SAH) upon income, as well as other relevant explanatory variables, one can see whether the term for income is estimated as positive and statistically significant. If it is, it is proven that absolute income positively affects the health of the individual, precisely in line with what the AIH dictates.

Equation 1 - Absolute Income Hypothesis

$$SAH_i = \alpha + \beta_1 X_i + \beta_2 income_i + \varepsilon$$

In Equation 1, the SAH of individual i is regressed upon X_i and income, where X_i is a vector of all other explanatory variables for health. If the AIH holds true, a regression would estimate β_2 as positive and statistically significant.

Regressing the model in Equation 1 is enough to prove the AIH but insufficient to also prove the RIH. To do the latter, a term for relative income must be included into the model. This study uses two different ways of creating such terms, one using objective measures of relative income and one using subjective ones. The specification of the two respective models will be explained below.

4.2 The two measures of relative income

The Relative Income Hypothesis (RIH) predicts a positive correlation between the relative income of the individual - compared to others - and the health of the individual. This study uses two different ways of capturing how the income of the individual relates to the income of others.

First of all, the stated income of all individuals is used to establish the average income for different groups, towards which the individual's own income can be compared. This way variables are created that show the mean and median income of the peers of the respondents. Although these variables do not technically offer information on the respondents' relative income they can be entered into regressions to find how respondents are affected by the income of others as a change of income of others yields a change of the individuals' relative income. This method establishes the relative income of the individual in objective terms and is hence called the objective measures of relative income.

The second way to establish relative income is to use the individual's estimation of how the own income relates to the income of others. This method directly establishes the relative income of the individual which can then be entered into regressions as an explanatory variable of Self-Assessed Health (SAH). Since this method establishes a term for relative income through subjective perception rather than stated income, it is called the subjective measure of relative income.

Using two different methods for establishing the relative income of the individual has benefits. Using objective measures is good as they are easily compared across individuals and their measuring is not subject to individual assessment. However, also using subjective measures makes the findings of the study robust against possible misreports of the own income. Further it also captures how the individual believes his or her relative income to be. There are several ways in which the subjective feeling of being richer or poorer than one's peers matters, and the individual's own perception regarding his or her income may very well be at least as important for determining health as the relative income in objective terms is (Singh-Manoux et al. 2005).

In the study on Socioeconomic Status (SES) and health outcome by Singh-Manoux et al (2005), a causal relationship from both subjective and objective SES to health status is found. This causal relationship holds true for both measures independently. However, when both measures are entered into a regression simultaneously, only the subjective measure of SES is found to be significant. This implicates that subjective SES is a better determinant of health status than objective SES. The authors believe this to be either because of psychological factors that relates the well-being of subjective SES to better health, or because of underlying personality characteristics that affect both the tendency to over-/underestimate the own SES and health status.

4.3 Model specification for the objective model

To test the RIH using objective measures of relative income a term for the income of a reference group is introduced. An example is shown in Equation 2. This reference group is a group which the individual belongs to and is expected to compare his or her income with. This reference group could theoretically be any group as long as the individual uses it for comparison or share or compete for resources with this group. Which reference groups are used in this study will be explained in the next section.

A model for testing the RIH using objective measurements of relative income is given by Equation 2, where X_i is a vector for all other explanatory variables for SAH.

Equation 2 – Relative Income Hypothesis (objective measures)

$$SAH_i = \alpha + \beta_1 X_i + \beta_2 income_i + \beta_3 reference\ group\ income_i + \varepsilon$$

When estimating Equation 2, the coefficient β_3 show how the income of a reference group affects the health of the individual. The RIH predicts a negative relationship, so that when the income of the reference group increases (i.e. the relative income of the individual decreases) the individual's health tend to be lower. By observing the direction and statistical significance of β_3, the RIH can be proved or rejected.

4.4 Creating reference groups

The importance of choosing the right reference group in order to accurately capture the effect of relative income has already been discussed. One practice used by Karlsson et al (2010) and McBride (2001) is to create reference groups based upon age. This study however uses geographical units for this purpose, in line with the discussions in section 3.2. This has the advantage of capturing the effects from the "price changes" mechanism (Miller & Paxson, 2006). It also seems reasonable that people compare their own income to that of their local geographic area, thus capturing the effects from psychological factors. Following the results and recommendations from Jones et al (2004), this study uses the most local level of geographical units available in the dataset: districts. The variable for reference group income is hence created as the mean income for the respondent's district. Due to the skewed income distribution of South Africa it is however not certain that mean income is a reasonable reference point as it may be strongly affected by the income of the few wealthiest individuals. For this reason another variable is also created: the median income of the respondent's district.

Although theoretically reasonable, there is a practical problem with only using a reference group that is already accounted for in estimation as an explanatory variable (for further discussion see Miller & Paxson, 2006). Assume that people in the local district constitute the primary reference point for the respondents. Then the RIH predicts a negative correlation between the health of the respondent and the average income of the own district holding individual income constant.

However, it seems reasonable to assume that there are also district-specific characteristics that affect health through other channels than income. This could be the result of district-specific health care policies, distance to health facilities, general attitude towards health behaviour or anything else that is unaccounted for in a regression. To account for this, "district" should enter into the model as an explanatory variable. This means that both

"district" and "reference group income" would be included as two separate variables in the same regression. This however creates a potential problem with multicollinearity as the district is both directly accounted for in the model and used to create the term reference group income that is also accounted for. Avoiding multicollinearity in the term for income of the reference group is vital to the study as a whole, as such a problem risks diminishing the statistical significance of the term, thereby potentially cause the Relative Income Hypothesis (RIH) to be wrongly rejected. Clearly, using an explanatory variable (e.g. "district") as a base for the so crucial reference group is problematic.

The solution used to overcome this problem is to only include district average (and median) income in the model, not using "district" as an explanatory variable. Since data is also available for "province", this can be used in regressions instead. This solution is in practice also used by McBride (2001). Although dropping the "districts" variable from the model eliminates the multicollinearity problem, it has the unfortunate drawback of the model failing to take into account a potentially important determinant of Self-Assessed Health. As some of the information available in the "district" variable is already captured in "province" (the greater geographical unit) the cost of dropping "district" should however be limited.

4.5 Model specification for the subjective model

An alternative way of testing the Relative Income Hypothesis (RIH) is to use the respondents (subjective) estimation of how the own income relates to the income of others. This method can be used by regressing the model given by Equation 3, where X_i is a vector of all other explanatory variables for health.

Equation 3 – Relative Income Hypothesis (subjective measures)

$$SAH_i = \alpha + \beta_1 X_i + \beta_2 income_i + \beta_3 perceived\ relative\ income_i + \varepsilon$$

When using the model given by Equation 3, the validity of the RIH is determined by the sign and statistical significance of β_3. The RIH predicts β_3 to be positive, indicating that higher relative income will be correlated to better health.

The study uses two different variables for perceived relative income. The first one compares the own income to that of others in the respondent's village/suburb. The second one compares the own wealth against the wealth of all other South Africans.

Unlike the objective measures of relative income, this measure only captures what individuals believe their own relative income to be. Because of this it is vulnerable to the respondent's own knowledge regarding the income of others. The upside is that this subjective measure can capture the feeling of being relatively rich or poor. Since the individual's perception of the own relative income may matter more than the actual relative income, these measures are valuable (Miller & Paxson, 2006).

Another difference towards the methods using objective measures of health is that there is no direct need of making assumptions regarding which group the respondents use for reference, since this is defined by the way the question was phrased in the interview.

4.6 Estimation method

The estimator used for this study is the *ordered probit regression*. Since the dependent variable is given on an ordinal scale (i.e. can only take a set amount of values), Ordinary Least Square (OLS) would yield biased estimations. Further, as the number of possible values is greater than 2, a multi-response model is needed, rather than a binary one. Given the ordered nature of the data (SAH=6 indicates better health than SAH=5, SAH=5 indicates better health than SAH=4 and so on), ordered probit regression and ordered logistic regression remains as the only two viable options (Verbeek, 2012). Just as McBride (2001), the ordered probit regression is chosen as it does not assume that an SAH increase from e.g. 1 to 2 is the same as from 3 to 4. The use of the ordered probit regression is further motivated by the large sample size as the assumption of normal distribution of the error terms is likely to be met (Verbeek, 2012). The reader should note that one needs to be cautious when interpreting the coefficients in ordered probit regressions. Whereas the coefficients in for instance the Ordinary Least Squares (OLS) method explain how much the dependent variable would be affected from a change to an explanatory variable, the ordered probit regression method allows no similar straightforward interpretation.

An alternative estimation method would be to dichotomize the data into just two categories: good and bad health. However, Karlsson et al (2010) states that although it is common to dichotomize Self-Assessed Health into a binary variable, this leads to less variability of the data and makes results vulnerable to the choice of cut-off points. Following the discussion above, ordered probit regression appears as the best choice for the study even though its limitations of interpretation.

In all models, the error terms are clustered at the household-level. The approach is due to that there are potential household-characteristics that will affect variation of the individuals' health status.

4.7 Further method considerations

The setup of the general tests for the Relative Income Hypothesis (RIH) should now be clear. This study however also employs some other tests regarding the connection between relative income and health. This section will briefly outline the respective reasons for these.

4.7.1 Does the RIH differ across income groups?

Previous studies have found that the size of an individual's income may affect whether the RIH holds true or not (Subramanian et al, 2001; McBride, 2001). Relative income may in other words affect income groups differently. As an example, it may be possible that relative income does not matter at all for poor people, whereas it matters substantially for wealthy individuals. It is important to know if different income groups are indeed affected to different extent for two reasons. First, research attempting to test the RIH could unfairly reject the hypothesis even for income groups where it holds true if groups where it does not would be included in the studies. Secondly, if the RIH holds true any policy choices stemming from this should be targeted towards those (income) groups where they would have the biggest impact. For these reasons, this study attempt to find out if the relevance of the RIH differ across income groups.

In order to test this, this study will run the regressions shown above on sub-groups of the respondents based upon income. The idea is to choose an income level and then run regressions to test the RIH on only the part of the population that has a household income above/below this level. This way it is possible to test if the RIH can be found for different sub-groups, not just the population as a whole. In this study income level quintiles are used to decide the income level cut-off points. As any differences between sub-groups and the general population is expected to be largest for the lowest and highest income groups, no regressions are performed for the three middle income quintiles. This means that the RIH is tested on sub-populations consisting of only the poorest and riches 20% of the population respectively.

4.7.2 Intertemporal correlation between relative income and health

Following the discussion in section 3.4, testing if there is an intertemporal correlation between health and relative income is interesting for understanding the mechanics through which they are related. Because of this two variables for relative income in other time periods than the current are included in this study. The first one is the respondents' estimation about the relative economic status of their household at the time they were 15 years old. The second one shows the respondents' estimation of the same status five year into the future. Although the latter is just, at best, qualified guesses by the respondents, this variable may still hold important information regarding health, regardless of what the future income actually turns out to be. This can be because predicted income may influence decisions regarding investment in the own health (Grossman, 1972). It can also be an effect of psychological factors, since perceived social mobility has been shown to matter more to subjective well-being (SWB) than actual social mobility (Fischer, 2009). The paper of Røysamb et al. (2003) on SWB and health finds that the subjective well-being (SWB) show a close relationship to perceived health.

4.7.3 Alternative measures of absolute income

As explained in section 4.1 it is crucial that a model for testing whether relative income affects health also includes a term for absolute income. Otherwise the impact of relative income may be overestimated as it not just captures the impact of relative but also of absolute income. Having an accurate measure of absolute income is therefore of paramount importance for this study. Although the measure used is in theory a good one, it is wise to also include some other measure to test if it works in practice.

In this study an alternative measure of absolute income is used as a robustness check of the results. This alternative measure is expenditure. Several researchers argue that is a good proxy for income when studying newly industrialized countries (Ravillon, 2003). First it takes into account the value of home production which in newly industrialized countries is substantial (Hjortsberg, 2003). Secondly it may fluctuate less than income over the year (Deaton & Zaidi, 2002). Given these reasons, alternative tests of the RIH will be performed using (absolute) expenditures in place of (absolute) income. If these tests come to the same conclusions as the ones using income then these measures likely capture absolute income accurately. If not it raises questions regarding if the terms for relative income is actually measuring what they are meant to.

5 - Data

This chapter will provide a brief description of the dataset used in this study. The first section will explain the origin of the data and how it has been collected. The following sections will discuss the creation of the variables used in the regressions. The chapter concludes with some descriptive statistics for the relevant variables, aimed at aiding the reader in understanding the South African setting. A list of all variables used in regressions is available in Appendix A.

5.1 About the data

The data used in this study is based on the National Income Dynamics Studies (NIDS), a national representative data over South Africa based on surveys conducted in 2012. The NIDS is a government funded project by the South Africa Labour and Development Research Unit (SALRU) and is operationalized by the University Of Cape Town´s School Of Economics. The dataset consist of both household and individual data over multiple social and economic factors with the purpose of getting a better understanding of poverty in South Africa (NIDS, 2015).

The data was collected by randomly selecting households and then interviewing all the residents within these households. The dataset contain information about 22,481 individuals that NIDS tried to reach. However, not all of these participated in a successful interview. Some individuals could not be reached or refused to take part in the interview. A total of 18,214 individuals from 7947 households were successfully interviewed and provided answers for all relevant questions. These are the respondents that all regressions are based upon.

In total, 4267 individuals did not provide sufficient answers to be a part of this study. This is around 19% of the potential respondents. This is not ideal as this could bias the results should any personal characteristics affect the chance of not participating in the interview. However, as an adequate explanation for why the interview failed was only recorded for few of these, further investigation of how the non-response rate affects the results is not possible.

The fact that households, rather than individual respondents, were randomly selected is a good way of ensuring unbiased data for households. However, it does not ensure an unbiased selection of the respondents. If individual characteristics affect the probability of

living in households of certain sizes, this way of selecting respondents may lead to a skewed sample of individuals in comparison to the actual South African population. If this is the case, this may slightly bias the findings of any study using the data material. However, the data in general reassembles official statistics reasonably well (compared with CIA World Factbook [a], 2015), so there is no need to further question the collection methods.

5.2 Health

This study uses Self-Assessed Health (SAH) as the main variable for health. This is generally considered a good way of measuring health status, and has gradually replaced mortality as the main proxy for health in the literature on the Relative Income Hypothesis (Wagstaff & van Doorslaer, 2000). The use of SAH allows overall health to be measured in one comprehensive variable, instead of having to weight different objective measures (such as the presence of certain diseases) together. Unlike many objective health measures SAH can also capture non-biomedical factors that are critical to an accurate health assessment (Eriksson et al, 2001). Furthermore, SAH has been shown to be a good predictor of both mortality and morbidity (Subramanian et al, 2001).

The problem with using Self-Assessed Health (SAH) is that people may differ in their reporting behaviour, meaning that the same health status can translate into different grading for different individuals. This is problematic as regressions attempting to find the correlation between health and income may instead end up catching differences in reporting behaviour. In their study of differences in health-related reporting behaviour, Subramanian et al (2001) find evidence that reporting behaviour can vary, not just across geographical units but also between income groups. The ability to assess the own health is influenced by the information and knowledge available to the respondents so that social context and comparison group will affect SAH outcomes. Consequently there might be heterogeneity in reporting SAH across subgroups, where the poor overestimate their health, as a result of different reference groups (Lindeboom and Van Doorslaer, 2004). This problem has been found to be especially large in Sub-Saharan Africa (Roussow, 2015).

The validity of SAH in the NIDS dataset is tested in the study by Laura Roussow (2015) where a comparison is performed between the reported SAH against objective health measurement in South Africa. The study finds that the poor South Africans tend to overestimate their health status and consequently underestimate their need for health care.

In this study the correlation between income and health may consequently be underestimated. However, SAH is still used to measure health as it remains a common practice in the academic literature and in lack of other more accurate objective measures of health.

5.3 Income

The dataset offers data on individual income divided into several different sources, such as income from the labour market, investments and grants. Although it is hence possible to use total individual income as the key variable, this is not the case in this study. Instead total household income per resident is used as the explanatory variable for income in the regressions. This is because the South African setting is substantially different from the countries where the absolute income hypothesis was first conducted in.

There are multiple reasons that household income is used instead of individual income. First of all, a large part of the respondents do not work for take home pay. The type of job done by people in the data is displayed in Table 5.1. It is likely that specialisation takes place within households so that some individuals would receive all the monetary income whilst others get none, regardless of how many hours they actually work. Due to this, resource pooling within families may be substantial. The existence of resource pooling in South Africa has been empirically tested and found to be true in a study by Case (2004). It has also been suggested that resource pooling is particularly likely when the health of one household member is at stake (Hjortsberg, 2003). Because of this individual income poorly reflects the actual amount of money the individual has at his or her disposal. Using household income instead of individual income is a way of overcoming this problem.

Second, using household income per resident rather than just individual income also takes into account grants targeting families. Grants such as child support would hence be omitted from the data if individual income was used, even should it contribute significantly to the wealth of the individual.

Third, pooling the income of several individuals may be a way of reducing the impact of temporary fluctuations to income. Since the data only contains the income of individuals for the last month, there is a substantial risk that this income does not accurately represent the usual amount earned by the individual. In lack of more data, this problem is impossible to completely overcome. However, the chance is smaller that all individuals within a household

experience the same magnitude of income fluctuations in the same month. Therefore pooling individual income into household income can marginally reduce the impact of these temporary fluctuations. This effect should not be overestimated though, as a household may rely on a single provider, hence be subject to the same fluctuations to a single individual's income.

Last and perhaps most importantly, using household income makes it possible to use both objective and subjective measures of relative income and see how they differ. The respondents were asked about the size of their relative household income. A similar question for relative individual income was not posed. Objective measures of relative income could just as easily have been derived if individual data had been used. However, household income ought to be used when dealing with subjective measures as this is how the questions were phrased in the survey.

Table 5.1

Last year's occupation	Freq.	Percent
Working for pay	4758	26.00
Self-employed	462	2.52
Working on own plot or looking after livestock	20	0.11
Helping another family member with their business without pay	33	0.18
Full-time scholar student at school, university, college or other educational institution	3247	17.74
Homemaker (looking after children/others / home)	1048	5.73
Long term sick or disabled	546	2.98
Retired	1315	7.18
Unemployment and actively searching for a job	3636	19.87
Unemployment but not actively searching for a job	3229	17.64
Refused	4	0.02
Don't know	5	0.03

Table 5.1 – Frequency table over the primary occupation held by the respondents over the previous year.

In line with the discussion above the variable for income used in the study is the total household income divided by the number of adult residents in the household. This total household income consists of income from different kind of sources, such as labour market

income and government grants (illustrated by Figure B.1 in Appendix B). Household income is used in its logarithmic form to account for the non-linear relationship between income and health. Since the data material used does not contain any zero values for income, this transformation does not alter the results.

Above it has been argued that for the purpose of this study, household income is preferable to individual income. However, income is not a perfect measure of economic ability regardless of the form it is measured in. In theory wealth is just as relevant in explaining health status as income. Further, as mentioned in section 4.7.3, consumption also has some desirable features as a proxy for income. For a more thorough discussion on wealth, income and consumption as measures the interested reader is referred to Deaton & Zaidi (2002).

5.4 Other variables

The perhaps most important determinant of health is age. Age is expected to be negatively correlated with Self-Assessed Health (SAH) as older individuals generally are in worse health condition than younger ones. However, this relationship may be non-linear, so that the marginal impact of age on health increases with age. This negative non-linear relationship is illustrated by Figure 5.1 of the data. This is accounted for in the regressions by including two explanatory variables for age, one linear and one quadratic.

Figure 5.1

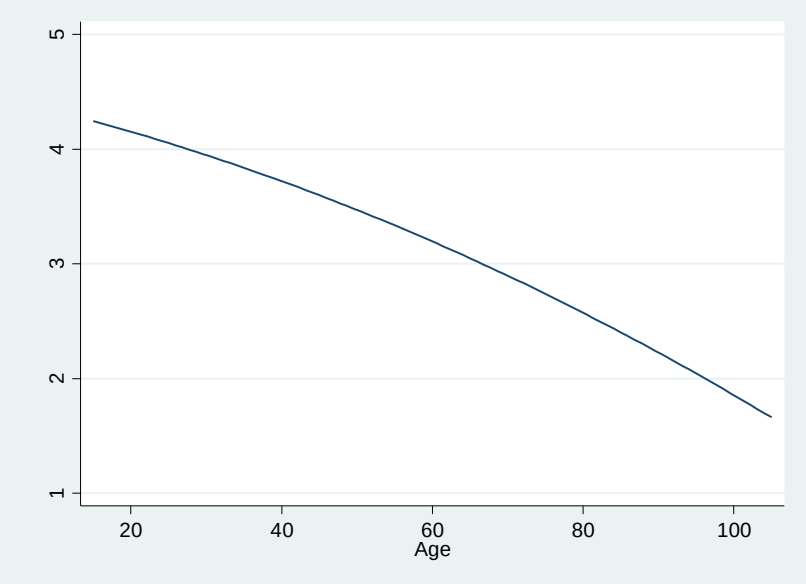

Figure 5.1 - Fitted values for Self-Assessed Health (SAH) as a function of age

Not all individuals within the South African education system follow a straight line in their education. An example is individuals who study technical skills at an equivalent level as theoretical education. To make such different education types comparable, the variable for the highest level of education of the individual in the dataset is transformed into four different education groups to obtain accurate estimations of its effect on health and to ease for interpretation. These groups are based upon the level of schooling which the last completed grade was part of. The four different education groups are: No education, Primary school, Secondary school and Higher education. These education groups are included the models as dummies to control for their impact on self-assessed health.

Studies have shown that women tend to report worse health than men (Månsdotter et al, 2004). Because of this a dummy variable for gender is included in the models. The type of household setting (urban, traditional and farms) are also included as regressions show this to be significantly correlated with health. For the same reason regional differences are also controlled for by using a dummy variable for "province". A term that is included without being analysed is "month of interview". It is included to avoid biasing the results due to seasonal variations as income in South Africa appear to change substantially over the year.

Similarly to the tests using variables for income, a variable for expenditure is also created as household expenditures per residents. This variable is used when performing the robustness checks, where expenditure is used as a proxy for income.

For a complete description of the variables used in this study readers are referred to the Table A.1 in Appendix A.

5.5 Descriptive data

As a last step before presenting the results some descriptive data of South Africa will be presented. This section aim to highlight some important aspects of the data. It also aims to help the reader understand that South Africa in some aspects differ importantly from the developed countries where the Relative Income Hypothesis (RIH) has predominantly been tested.

Self-Assessed Health (SAH) is the health status estimated by the individuals themselves and it is measured on a 5 degree scale where 1 is the lowest health and 5 the highest according to Table 5.2. The data show that almost 9 out of 10 South Africans rate their health to be at least good.

Table 5.2

Self-Assessed Health	Frequency	Percentage
1- Poor	600	3.28
2- Fair	1612	8.81
3- Good	5226	28.55
4- Very good	5351	29.24
5- Excellent	5514	30.13

Table 5.2 – *Frequency table over Self-Assessed Health*

The data over South Africa shows vast income differences within the country which is illustrated by Table 5.3. When asked to classify the own household income on a six grade scale along with all other households in South Africa, 88,3% of the respondents state that they belong among the three lowest grades. This is further emphasized on the local level in Table 5.4 where about half of the respondents state a wealth below the village average and only 9,4% of the respondents believe themselves to be “above average” or higher.

Table 5.3

Relative Income (nation)	Frequency	Percentage
One (Poorest)	2628	14.36
Two	6884	37.61
Three	6648	36.32
Four	1762	9.63
Five	318	1.74
Six (Richest)	63	0.34

Table 5.3 – *Respondents' perceived relative income (nation)*

Table 5.4

Relative Income (village)	Frequency	Percentage
1 - Much below average	3764	20.56
2 - Below average	5625	30.73
3 – Average	7203	39.35
4 - Above average	1305	7.13
5 - Much above average	406	2.22

Table 5.4 – *Respondents' perceived relative income (village)*

The data shows that a vast majority consider themselves to have average income or lower compared to the own village/suburb. This is intriguing. This may be the result of a few wealthy individuals raising the average income so substantially that most respondents consider their earnings below this. It may however also be the result of lack of knowledge or a general tendency to underestimate the own income in comparison to others. Unfortunately one can only speculate as to how the data should be interpreted. Clearly even straightforward questions can give rise to puzzling answers in surveys.

South Africa has a young population. Interestingly enough, many respondents are not much older than 15 years - the minimum age to take part of the study. The frequency distribution for age is presented by Figure B.2 in Appendix B.

The huge income inequality discussed in section 2.1 is also evident in the data sample. The absolute income of the respondents ranges widely from only 12 Rand/month to 85000 Rand/month. There is also a big difference between mean and median income in the dataset, indicating that the few richest respondents in the sample has income so high that it substantially raises the mean for the whole sample. The income distribution is shown in Table 5.5. It is also presented graphically by the Lorenz curve in Figure B.3 in Appendix B.

Table 5.5

	Income
Minimum	12
20%	365
40%	597
Median	750
60%	952
80%	1863
Maximum	85000
Mean	1572

Table 5.5 - *Distribution of income across different income percentiles in the dataset. All income is given in South African rand and calculated as household income per resident and month. The percentages indicate the share of the population that has a lower income than the given value. Minimum and maximum show the lowest and highest values in the dataset respectively.*

6 - Results

This chapter will present the results of the main regressions on the Absolute Income Hypothesis (AIH), the Relative Income Hypothesis (RIH) using objective measures and the Relative Income Hypothesis (RIH) using subjective measures. The results of the main regressions are presented in six separate models in Table 6.1. After this the results from the tests of the RIH over income groups are presented in Table 6.2. Thereafter the results from the time effects of past and expected future income are presented in Table 6.3. To conclude this chapter the robustness checks with expenses are presented in Table 6.4.

6.1 Main Results

First a basic estimation of the Absolute Income Hypothesis (AIH) on the Self-Assessed Health (SAH) is performed which is presented in model 1 in Table 6.1. This version of testing the AIH includes the main explanatory variable Log Income, as well as the basic characteristics of the individual with age, gender and population group. The results from model 1 indicate that income have a positive significant relationship to SAH. This supports the AIH.

The test is then extended by regressing model 2, which includes additional explanatory variables to obtain a more accurate estimation of the AIH. In model 2 the positive significant relationship between income and SAH remains, providing more evidence to the AIH. The coefficient for the income variable remains positive and significant in all the models, supporting the AIH in that higher household income is associated with better health.

The tests for the Relative Income Hypothesis (RIH) are performed and presented as models 3-6 in Table 6.1. These models are based on the full model specification of the RIH with the same explanatory variables included as in model 2, and they include a different variable of relative income each. The first test of the RIH with an objective measure using the reference group mean variable is presented in model 3. The results from this estimation show a positive but insignificant coefficient for reference income (mean), the results finds no evidence for the RIH. The results of model 4 where the reference group median is measuring the RIH, show a positive but insignificant coefficient. Hence model 4 shows no evidence of the RIH being true. In conclusion, both models that use objective measures of the reference group income finds no support for the RIH.

The results for the estimation of the RIH using subjective measurements are presented in model 5 and model 6 in Table 6.1. In model 5 the variable of relative income (village), show a

positive significant coefficient. The results imply that, when the individual becomes relative richer to the village, holding the individuals income constant, the individuals health improves. Consequently the results in model 5 support the RIH. In model 6 the subjective measurement of relative income is a ranking towards South Africa as a whole. The coefficient for this variable show a positive significant coefficient, which further supports the RIH. The models using subjective measures of relative income hence both finds support for the RIH. The individual's relative income, holding absolute income constant, is positively related with better health, regardless if income is compared to the village- or nation-level.

The results from the additional explanatory variables show that there are many factors that influence the individuals reported health. Age has through the models shown a negative significant coefficient, indicating that the health status decrease over time for individuals. The variable Age^2 show a negative significant coefficient in model 1, but is insignificant for all other models. It can hence not be proved that the health impact from age increases with rising age. Women seem to report worse health which is indicated by the negative significant coefficient of the Female variable. The Province variable remains significant in all models, indicating that it has explanatory value of assessing the health status of the individual. Education has, as suspected, a large positive effect on the individuals' health and is therefore relevant to include in the models when estimating health. The martial status, also show a positive significant coefficient: individuals that are married or living with a partner tend to report better health. These additional explanatory variables will not further be presented or analysed as they are only included to obtain more accurate estimations of incomes effect on health.

Table 6.1

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	Absolute Income Hypothesis	Absolute Income Hypothesis	Relative Income Hypothesis - Objective	Relative Income Hypothesis - Objective	Relative Income Hypothesis - Subjective	Relative Income Hypothesis - Subjective
Log Income	0.0465635 ***	0.0377888 ***	0.0367918 ***	0.0368087 ***	0.0311245 **	0.0246541 *
	(0.0110201)	(0.0123394)	(0.0123363)	(0.0123633)	(0.012784)	(0.0126759)
Age	-0.0212526 ***	-0.025416 ***	-0.025462 ***	-0.0254778 ***	-0.0250416 ***	-0.0244564 ***
	(0.0022232)	(0.0024554)	(0.0024561)	(0.0024556)	(0.0024568)	(0.0024517)
Age^2	-8.71e-06	0.0000236	0.000024	0.0000242	0.0000197	0.0000152
	(0.0000249)	(0.0000267)	(0.0000267)	(0.0000267)	(0.0000267)	(0.0000267)
Female	-0.1117166 ***	-0.1117935 ***	-0.1116897 ***	-0.1116281 ***	-0.113301 ***	-0.1149622 ***
	(0.0159224)	(0.0160476)	(0.0160432)	(0.0160415)	(0.0160546)	(0.0160311)
Education Level	0.1403412 ***	0.1445517 ***	0.1441759 ***	0.1441579 ***	0.1420632 ***	0.138758 ***
	(0.0106758)	(0.0107539)	(0.0107341)	(0.0107399)	(0.0107265)	(0.0107797)
Reference income (mean)			0.0345983			
			(0.0546275)			
Reference income (median)				0.0353007		
				(0.0511488)		
Relative Income (village)					0.0323601 ***	
					(0.0114312)	
Relative Income (nation)						0.0566145 ***
						(0.0109748)
	N=18214	N=18214	N=18214	N=18214	N=18214	N=18214

Table 6.1 - Regression outputs testing the Absolute Income Hypothesis (AIH) and Relative Income Hypothesis (RIH). The dependent variable is Self-Assessed Health (SAH). Models 2-6 are estimated using dummies for "province", "population group", "married", "residential type" and "month of interview". Standard errors are clustered at the household-level. * p <0.1, ** p<0.05, ***p<0.001

6.2 Relative income hypothesis for different income groups

The tests of the income groups show different results for the subsamples, as shown in Table 6.2. For the poorest population of South Africa in model 7, the RIH is not found when it is measured in objective terms. However when the subjective measure is used in model 8 the results show a positive significant coefficient which indicate support for the RIH. The test for the rich subsample, finds in model 9 a positive significant coefficient for the objective measure which is contradicting the RIH. Support for the RIH is also not found for the rich subsample when the subjective measure is used in model 10. To conclude, the results from the subsample tests find that the RIH is only found for the subjective measure for the poorest individuals in South Africa. Further investigation of the results for model 7-10 show that (absolute) income only have a significant positive relationship to SAH in model 9 & 10,

indicating that support for Absolute Income Hypothesis can only be found for the rich subsample.

Table 6.2 only show the results from one measure of objective and subjective income each for the poor and rich income groups respectively. Tests using the remaining 4 combination of variables and income groups were also performed. However, as the results were similar in terms of significance and sign of the term as the findings in model 7-10, the results of these tests are not presented.

Table 6.2

Variable	Model 7 RIH Bottom 20% Objective	Model 8 RIH Bottom 20% Subjective	Model 9 RIH Top 20% Objective	Model 10 RIH Top 20% Subjective
Log Income	0.0179445	0.0128939	0.104044 ***	0.098429 ***
	(0.0781419)	(0.0782598)	(0.0361604)	(0.0369655)
Age	-0.0287762 ***	-0.0283336 ***	-0.0123532 *	-0.0119103 *
	(0.0055292)	(0.0055226)	(0.0064581)	(0.006446)
Age^2	0.0000514	0.0000448	-0.0001119	-0.0001158
	(0.0000624)	(0.0000622)	(0.0000715)	(0.0000714)
Female	-0.0907999 **	-0.0945582 **	-0.1391417 ***	-0.141034 ***
	(0.038342)	(0.038299)	(0.0350804)	(0.0350632)
Education Level	0.11621 ***	0.1110299 ***	0.163969 ***	0.1617988 ***
	(0.0229357)	(0.022846)	(0.0304544)	(0.0304514)
Reference income (median)	-0.1590317		0.0898898	
	(0.1281946)		(0.0954889)	
Relative Income (village)		0.0643823 **		0.025348
		(0.0257729)		(0.0261414)
	N=3642	N=3642	N=3614	N=3614

Table 6.2 - Regression outputs testing the Relative Income Hypothesis (RIH) for the poorest and richest 20% of the population respectively. The dependent variable is Self-Assessed Health (SAH). All models are estimated using dummies for "province", "population group", "married", "residential type" and "month of interview". Standard errors are clustered at the household-level. * p <0.1, ** p<0.05, ***p<0.001

6.3 Intertemporal correlation

As pointed out in sections 3.4 and 4.7.2, health may be affected by more than just contemporaneous income. To test this, regressions are performed where current relative income (as compared to the nation) is replaced by relative income for when the respondent was 15 and the respondent's expected income 5 year from the time of interview

respectively. In both these cases significance is found for these terms, meaning that the RIH can also be valid for past and future income. However, even though a statistically significant positive correlation is found, it is in both cases smaller than the correlation between SAH and current relative income. It is therefore not clear if past and future income explains SAH outcomes except through explaining current income. To see if past and future income adds further explanatory value to the model, they are both included in a new regression together with current income. The results of these regressions is shown in Table 6.3.

Table 6.3

Variable	Model 11 RIH Subjective Past Income	Model 12 RIH Subjective Future Income
Log Income	0.0246413 *	0.023077 *
	(0.0126716)	(0.0126606)
Age	-0.0244723 ***	-0.0241732 ***
	(0.0024519)	(0.0024505)
Age^2	0.0000153	0.0000158
	(0.0000267)	(0.0000266)
Female	-0.115074 ***	-0.1127937 ***
	(0.0160286)	(0.0160481)
Education Level	0.1383704 ***	0.1348298 ***
	(0.0108068)	(0.0107977)
Relative Income (nation)	0.0532317 ***	0.0467434 ***
	(0.0113864)	(0.0109867)
Relative Income (nation) when 15	0.0061593	
	(0.0075511)	
Relative Income (nation) in 5 years		0.0148587 ***
		(0.0021448)
	N=18214	N=18214

Table 6.3 - Regression outputs testing the Relative Income Hypothesis (RIH) using past and predicted future income. The dependent variable is Self-Assessed Health (SAH). Both models are estimated using dummies for "province", "population group", "married", "residential type" and "month of interview". Standard errors are clustered at the household-level. * p <0.1, ** p<0.05, ***p<0.001

Both model 11 and 12 show that current income is significantly correlated to Self-Assessed Health (SAH) even when a variable for past or future income is included. However, the regressions output of model 11 show that the income of the household when the

respondent was 15 adds no (or at least very limited) further explanatory value to SAH. By contrast, model 12 indicate that predicted income for the future is a very significant determinant of health, even when accounting for current income. Apparently the belief about one's own future income is a better predictor of current health than past income.

6.4 Robustness checks

Sections 3.1 and 4.7.3 have explained why it is crucial to the findings of this study that the term for relative income does not capture the size of absolute income. To check the robustness of this study's results this term is now substituted for a term on household expenditures per resident. The regressions from models 3-6 are then performed again, this time using expenditures as a proxy for income. For the models using objective measures of relative income the term for mean/median income in the district is also replaced by a term for mean/median expenditures in the district. The results of some of the regressions are given in Table 6.4.

Further robustness checks using the second version of the objective (mean) and subjective (nation) are also performed but as their results are similar and adds no further findings compared to the other measures they are not presented.

Table 6.4

Variable	Model 13	Model 14
	RIH Objective Expenditures	RIH Subjective Expenditures
Log Expenditure	0.0296779 **	0.0244773 *
	(0.0126106)	(0.0129982)
Age	-0.0253898 ***	-0.024941 ***
	0.0024559	(0.0024581)
Age^2	0.0000249	0.0000201
	(0.0000267)	(0.0000267)
Female	-0.1146331 ***	-0.1158497 ***
	(0.0160353)	(0.0160434)
Education Level	0.14526 ***	0.1429573 ***
	(0.0107218)	(0.010709)
Reference Expenditure (median)	0.0397841	
	(0.047884)	
Relative Income (village)		0.033992 ***
		(0.0113365)
	N=18214	N=18214

Table 6.4 - Regression outputs testing the Relative Income Hypothesis (RIH) using expenditures as a proxy for income. The dependent variable is Self-Assessed Health (SAH). Both models are estimated using dummies for "province", "population group", "married", "residential type" and "month of interview". Standard errors are clustered at the household-level. * p <0.1, ** p<0.05, ***p<0.001

Comparing the results in Table 6.4 with Table 6.1 shows that the findings does not change much depending on whether household income or household expenditure is used as a measure of income. The RIH is still rejected when income is measured in objective terms and confirmed when subjective measures are used. A simple correlation test between income and expenditure confirm that they largely explain the same thing, with the test showing a 78.19% correlation. To further test if a measure of expenditures can add to the study both measures are simultaneously included into the same regression. When this is done only the terms for income remain statistically significant which can be seen as an indication that income is a better predictor of SAH than expenditure.

7 - Analysis & Discussion

The result from the main regressions find empirical evidence for the Absolute Income Hypothesis (AIH) in all but one model. These results were expected as multiple studies over AIH have found a positive relationship between income and health all over the world (e.g. Karlsson et al, 2010; Wagstaff & van Doorslaer, 2000; Case, 2004 and Gerdtham & Johannesson, 2004). Furthermore as this study was conducted over South Africa which is suffering from large income inequality, poor health and an unequal health care system, the effect of absolute income on health was expected to substantial.

The results for the Relative Income Hypothesis (RIH) are somewhat inconclusive. In a way, this too was expected as earlier studies have found mixed evidence on the topic. This study finds that whether support for the RIH is found or not is largely affected by the types of measure used to capture relative income. When using objective measures the RIH is not supported. When subjective measures is used instead the RIH is however supported. Why these two types of measures yields such different results is an important question for this study.

The different findings for objective and subjective measures can both reflect that the models partly measures different mechanisms of RIH and/or the difficulties to accurately test the RIH. A potential threat when estimating the RIH is that it also captures part of the effect of absolute income. In lack of more theoretically sound measures of absolute income, it is impossible to know how severe this problem is in the study.

The tests for the two different objective measures of the RIH (mean and median) finds no significance in any of the models, therefore no support for the RIH is found by the objective measure. The coefficient for the objective measure in most of the models are positive even though the lack of significance limit the interpretation. This reversed relationship could reflect a true negative relationship between the relative income and health status, possibly explained by large spill-over effects from the local community. However, given the low level of taxation and policies on wealth distribution in South Africa this is hardly the full explanation. A perhaps more reasonable explanation is that absolute income effects are partly captured in the measure for relative income. Consequently the positive relationship between district average income-levels and Self Assessed Health (SAH) could simply reflect

that the individual lives in a wealthy district, which is likely to be determined by the individuals' income.

The reference groups used for deriving relative income (using objective measures) are in this study based upon districts. It is possible that districts are too large to accurately pick up the effects from relative income, should they exist. After all, the theory states that it is the reference group of the individual that matter for his or her health. If reference groups are in practice based upon smaller geographical units than districts, such as village, it is possible that the results would have been different. Use of another dataset would perhaps have been able to shine a light on this. Although the models using objective measures were unable to confirm the RIH, one cannot confidently say that they rejected it.

By contrast, the models using subjective measures of relative income finds support for the RIH. A higher ranking of the own wealth, compared to either the local village/suburb or the entire nation, is correlated with better health. This is in line with the theory of RIH and previous studies using subjective measures, such as Eibner & Evans (2005), Subramanian et al (2001) and Miller & Paxson (2006).

The models using subjective measures for relative income are likely to catch psychological effects as they are based upon the respondents' perception of the own income. Consequently even though the individual might not be relatively poor, the individuals' experience of feeling poor will affect the SAH. The results from the subjective measure are in line with the studies that the perception of being relative poor is strongly associated with psychological stress and negative health outcomes for the individual (Grönqvist et al, 2012; Eibner & Evans, 2005; Miller & Paxson, 2006).

It is paramount to the conclusions of this study that the regressions can fully distinguish between absolute and relative income. If they do not the RIH may be wrongly accepted or rejected. In lack of good ways to ensure that the variable for income truly does capture all absolute income effects, this study has used variables for expenditures as a proxy for income. The results of model 13-14 in Table 6.4 shows that the findings are at least moderately robust to different measures of income. In line with the discussion in section 4.7.3, this strengthens the reliability of the study. However, even though the findings have

been show robust to this alternative measure of income, there is still no way of proving the accuracy of the measures for absolute income.

The fact that household income is used in place of individual income is another source of uncertainty. This study has argued that household income per resident is the best measure of individual income available. However, as economic power is not equally distributed within households, future studies should focus more on the size of economic resources disposable to each respondent.

The different results from models using objective and subjective measures of relative income is an important finding of this study. It does show the importance for future studies of including both these measures, or at least address this issue. It also provides a possible explanation as to why the findings of earlier studies have differed so much. However it also raises a challenging problem: which type of measurement should be trusted? The subjective measures allow a study to capture the impact on health from psychological factors such as the belief about ones relative wealth or ones future income. If researchers believe these effects to be substantial subjective measures are preferable. If not then objective measures have the edge since it is easier to compare individuals within the sample. Naturally including both types of measures gives the advantages of learning from how the results differ.

This study has further attempted to find if there are any structural differences between different income groups regarding how relative income affects health. With only one out of four tests (subjective measures for the poorest 20%) finding statistical significance the results points towards this not being a major factor. This also goes against the evidence by the earlier study by McBride (2001) which found relative income effects to be of most importance for the wealthier groups. Naturally, rejecting differences between income groups could be a false conclusion: given that these tests are performed at only 20% of the sample, finding significance could be hard, even if differences existed. For instance the tests for the AIH, which always supported the theory when performed at the whole sample, suddenly became inconclusive when tests were performed at only 20% of the sample size. However, if the relationship between relative income and health looked vastly different between the groups, this would likely have showed more clearly in the results. To sum things up, the set-up of this study is either too basic in design or too small in sample size to either

confirm or reject structural differences between income groups. It does however imply that should such differences exist, they would be of limited importance.

This study finds that health is correlated to wealth in other time periods than the present. Both past wealth and predictions about future wealth have been found significant in explaining current health. However, the results from model 11 in Table 6.3 show that past income loses its explanatory value when current income is simultaneously included in a regression. Hence it is not proven that past wealth influences health, other than indirectly through its effect on current wealth. This finding is interesting as it may also say something about causality between health and income. If there is a strong causal relationship from income to health, then past income should be a great predictor of current health. The findings do not disprove this, but they indicate that past income is more relevant to explaining health status through its correlation to current income than through direct health effects.

It should also be mentioned that South Africa has a young population as shown in Figure B.2 in Appendix B. It is therefore possible that the past income effects have not yet had their full impact on health and that past income does not differ that much from current income. Perhaps the findings regarding past income and current health would have been different for an older population.

By contrast, predicted future wealth remains a significant determinant of current health, even when current wealth is simultaneously entered into a regression, as shown by Model 12 in Table 6.3. There are two likely explanations for this. First of all, health can affect earnings prospects, so that people consider their own current health state when asked to predict their future income. Secondly, the beliefs about one's own future income prospects may have direct impact upon current health, even if this does not reflect the actual future earnings. The latter would indicate that psychological factors play an important role in determining health status, and those individual characteristics regarding expectations of the future are crucial in explaining health. If this is true, it could of course lead to beliefs about the future being self-fulfilling.

A better understanding of how income differences affect the health status of the population is crucial for policy-makers and organisations interested in improving societal health. This

study has found strong evidence that the AIH holds true, even for newly industrialized countries. Although the evidence regarding the RIH is mixed it still provides insight into important considerations for the methodology for testing the theory. Hopefully this can help future studies reach a bit further in understanding the relationship between health and relative income.

The findings of this study differs considerably from some previous ones. However, as neither its methods nor the setting in South Africa is similar to previous studies any comparisons of findings need to be done with great caution. The connection between income, relative income and health is indeed a vastly complex topic.

8 - Conclusion

This study has attempted to find out if the Relative Income Hypothesis (RIH) is relevant in explaining health outcomes in South Africa. It has found that the result hinges on the types of measures used for relative income. When using objective measures of relative income, no support for the RIH is found. However, the RIH is supported when subjective measurements are used. One explanation for this is that psychological effects, such as the feeling of being poor, have a big impact on health status. The fact that belief about future income is highly correlated to current health strengthens this theory. Another possible, although somewhat less likely, explanation is that health status may be affecting the reporting behaviour of subjective wealth more than it affects actual income. However, the support for the RIH may also be the result of a failure to completely capture all effects of absolute income in the model.

The study also attempt to test if the relevance of relative income for health status may differ across income groups. Although it finds indications that differences may exist no systematic conclusions can be drawn.

Health status have also been found correlated with both past and predicted future wealth. It is however only predicted future income that remains significant when also controlling for current income. This may indicate that psychological factors are relevant in explaining current health status.

The study further finds that the Absolute Income Hypothesis (AIH) is relevant for explaining health outcomes for individuals in South Africa.

This study has tested the RIH for South Africa. Although findings are not transferable to other newly industrialised countries per se, it is likely that both findings and challenges are similar. The RIH should be further studied in newly industrialized countries as its impact and relevance may differ entirely from that of the developed world.

In lack of a complete understanding of the mechanisms through which the RIH works, future studies should combine both objective and subjective measures. They should also pay particular attention to ways of detangling relative income effects from absolute income effects. Ensuring accurate measures of absolute income is a crucial step. More focus on which resources are available to individual family members may be a way of achieving this.

9 - References

Benzeval, M. & Judge, K. (2001) Income and Health: the time dimension, *Social Science and Medicine*, vol 52(9), pp. 1371-1390

Blanco-Perez, C. (2012). *Rethinking the Relative Income Hypothesis*. DIW Berlin, The German Socio-Economic Panel (SOEP) (No. 501).

Case, Anna (2004) "Does Money Protect Health Status? Evidence from South African Pensions" The National Bureau of Economic Research Volume: *Perspectives on the Economics of Aging*

Clark A., Frijters P. & Shields M. (2008) Relative income, happiness, and utility: An explanation for the Easterlin paradox and other puzzles, *Journal of Economic Literature* vol 46(1), pp. 95-144.

Coovadia, H., Jewkes, R., Barron, P., Sanders, D., & McIntyre, D. (2009). The health and health system of South Africa: historical roots of current public health challenges. *The Lancet*, *374*(9692) pp. 817-834.

Deaton, A. & Zaidi, S. (2002) Guidelines for Constructing Consumption Aggregates for Welfare Analysis. *Living Standards Measurement Study Working Paper 135*. Washington DC: World Bank.

Eibner, C. & Evans W. (2005) Relative Deprivation, Poor Health Habits, and Mortality, *Journal of Human Resources*, vol 40(3), pp. 591-620

Ellison, G. (2002) Letting the Gini out of the bottle? Challenges facing the relative income hypothesis, *Social Science & Medicine* vol. 54(4) pp. 561-576.

Erdil, E., & Yetkiner, I. H. (2004). *A Panel Data Approach for Income-Health Causality* (No. FNU-47). Research unit Sustainability and Global Change, Hamburg University.

Eriksson, I., Undén, A. & Elofsson, S. (2001) Self-rated Health. Comparisons between three different measures. Results from a population study, *International Journal of Epidemiology*, vol 30, pp. 326-333

Fischer, J. A. (2009), "The Welfare Effects of Social Mobility", OECD Social, Employment and Migration Working Papers, No. 93, OECD Publishing accessed 2015-06-05
http://dx.doi.org/10.1787/221272634852

Fuchs, V. (2004) Reflections on the socio-economic correlates of health, *Journal of Health Economics*, vol 23, pp. 653-661

Gerdtham & Johannesson (2004) Absolute Income, Relative Income, Income Inequality and Mortality, *Journal of Human Resources*, vol 39(1), pp. 228-247

Gravelle, H (1998) How Much of the Relation Between Population Mortality and Unequal Distribution of Income is a Statistical Artefact?, *British Medical Journal*, vol 316, pp. 382-385

Grossman, M. (1972) On the Concept of Health Capital and the Demand for Health, *The Journal of Political Economy*, vol 80(2), pp. 223-255

Grönqvist, H., Johansson, P. & Niknami, S. (2012) Income Inequality and Health: lessons from a refugee residential assignment program, *Journal of Health Economics*, vol 31, pp. 617-629

Hjortsberg, C (2003) Why do the Sick Not Utilise Health Care? The Case of Zambia, *Health Economics*, vol 12, pp. 755-770

Jones K., Duncan, C. & Twigg L. (2004) Evaluating the absolute and relative income, hypotheses: an explanatory analysis of deaths in the Health and Lifestyle Survey, in Boyle P. (ed.) *The geography of health inequalities in the developed world*, Aldershot: Ashgate

Karlsson, M., Hansson, T., Lyttkens, C. & Neeson, G. (2010) Income Inequality and Health: Importance of a Cross-Country Perspective, *Social Science & Medicine*, vol 70, pp. 875-885

Kawachi, I., Subramanian, S. V., & Almeida-Filho, N. (2002) A glossary for health inequalities. *Journal of epidemiology and community health*, *56*(9), 647-652.

Lindeboom, M., & Van Doorslaer, E. (2004). Cut-point shift and index shift in self-reported health. *Journal of health economics*, *23*(6), 1083-1099.

McBride, M. (2001) Relative-income effects on subjective wellbeing in the cross-section, *Journal of Economic Behavior & Organization*, vol 45(3), pp. 251-278

Mellor, J. & Milyo, J. (2002) Income Inequality and Health Status in the United States: Evidence from the current population survey, *The Journal of Human Resources*, vol 37(3), pp. 510-539

Miller, D. & Paxson, C. (2006) Relative Income, Race and Mortality, *Journal of Health Economics*, vol 25, pp. 979-1003

Månsdotter, A., Lindholm L. & Öhman, A. (2004) Women, men and public health – how the choice of normative theory affects resource allocation, *Health Policy*, vol 69, pp. 351-364

Pritchett, L. & Summers, L. (1996) Wealthier is Healthier, *Journal of Human Resources*, vol 31(4), pp. 841-868

Ravallion, M. (2003) The Debate on Globalization, Poverty and Inequality: why measurement matters, *International Affairs*, vol 79(4), pp. 739-753

Roussow, L. (2015) Poor health reporting: Do poor South Africans underestimate their health needs? No. UNU-WIDER Research Paper

Røysamb, E., Tambs, K., Reichborn-Kjennerud, T., Neale, M. C., & Harris, J. R. (2003). Happiness and health: environmental and genetic contributions to the relationship between subjective well-being, perceived health, and somatic illness. *Journal of personality and social psychology*, *85*(6), 1136.

Singh-Manoux, A., Marmot M. & Adler, N. (2005) Does Subjective Social Status Predict Health and Change in Health Status Better then Objective Status?, *Psychosomatic Medicine*, vol 67, pp. 855-861

Subramanian, S., Kawachi, I. & Kennedy, K. (2001) Does the State You Live in Make a Difference? Multilevel analysis of self-rated health in the US, *Social Science & Medicine*, vol 53, pp. 9-19

Verbeek, M (2012) A Guide to Modern Econometrics, 4th Edition Wiley, New York

Wagstaff, A. & van Doorslaer, E. (2000) Income Inequality and Health: What Does the Literature Tell Us?, *Annual Review of Public Health*, vol 21(1), pp. 543-567

Wilkinson, Richard G (1999) Health, hierarchy, and social anxiety. *Annals of the New York Academy of Sciences 896.1* pp.48-63.

Data Sources

Brand South Africa (2015), accessed 2015-06-05 at:
http://www.southafrica.info/about/health/

CIA – The World Factbook [a] (2015) accessed 2015-06-05 at:
https://www.cia.gov/library/publications/the-world-factbook/geos/sf.html

CIA – The World Factbook [b] (2015) accessed 2015-08-19 at:
https://www.cia.gov/library/publications/the-world-factbook/rankorder/2172rank.html

NIDS (2015), What is NIDS, accessed 2015-06-05 at:
http://www.nids.uct.ac.za/about/what-is-nids

NIDS User Manual, accessed 2015-06-05 at:
http://www.nids.uct.ac.za/images/documents/wave3/NIDS-W3UserGuide-PublicRelease20150330_v1.5.pdf

US Embassy in South Africa, accessed 2015-06-05 at:
http://southafrica.usembassy.gov/root/pdfs/study_sa_profile_rev100630.pdf

WHO (2015) Statistical Profile: South Africa, accessed 2015-06-05 at:
http://www.who.int/gho/countries/zaf.pdf?ua=1

World Bank (2015), accessed 2015-08-19 at:
http://data.worldbank.org/indicator/SI.POV.GINI?order=wbapi_data_value_2011+wbapi_data_value&sort=asc

Appendix A – List of Variables

Table A.1

Name	Description	Coding
SAH*	Self-Assessed Health as stated by the respondents	1=Poor 2=Fair 3=Good 4=Very Good 5=Excellent
Log Income	Total logarithmic household income per month divided by the number of household residents	Min=12 ** Mean=1572 Max=85000
Reference Income (mean)	The district mean household income	Min=622 Mean=782 Max=4397
Reference Income (median)	The district median household income	Min=431 Mean=846 Max=2100
Relative Income (village)	The (subjective) relative income of the respondent compared to the local village/suburb	Much below average income Below average income Average income Above average income Much above average income
Relative Income (nation)	The (subjective) relative wealth of the respondent's household compared to the all of South Africa	One (Poorest) Two Three Four Five Six (Richest)
Relative Income (nation) when 15	The (subjective) relative wealth of the respondent's household compared to the all of South Africa when the respondent was 15 years old	Don't know One (Poorest) Two Three Four Five Six (Richest)
Relative Income (nation) in 5 years	The predicted (subjective) relative wealth of the respondent's household compared to the all of South Africa 5 years after the interview	Don't know One (Poorest) Two Three Four Five Six (Richest)
Log Expenditure	Logarithmic total household expenditures per month divided by the number of household residents	Min=26 ** Mean=1107 Max=65348
Age***	The age of the respondent	Min=15 Mean=38 Max=105
Female	A dummy-variable for the gender of the respondent	0=male 1=female

Population group	The population group that the respondent identifies him-/herself as. The distribution of respondents between the alternatives are given in percentages.	African 82,52 Coloured **** 13,66 Asian/Indian 1,02 White 2,80
Married	Relationship status of the individual.	Never married Widow/widower Divorced/separated Don't know Refused Married Living with partner
Residential type	The type of accommodation that the respondent live in. The distribution of respondents between the alternatives are given in percentages.	Traditional 43,67 Urban 47,82 Farms 8,51
Province	Which province the respondent live in	Western Cape Eastern Cape Northern Cape Free State KwaZulu-Natal North West Gauteng Mpumalanga Limpopo
Education Level *****	The level of completed education of the respondent. The distribution of respondents between the alternatives are given in percentages.	No schooling 10,96 Primary School 14,12 Secondary School 27,19 High School 47,73
Month of interview	Which month in 2012 that the interview was conducted	April May June July August September October November December

Table A.1 – List of variables used in regressions. For some variables the descriptive statistics are presented.

* The survey was originally coded in reversed order so that 1 corresponded to excellent health. This order was however changed in this study to make the findings on relative income easier to interpret.

** These figures are given in their non-logarithmic form to aid the reader

*** Household members below the age of 15 were also interviewed, but they were given slightly different questions, which were presented in a separate dataset. Because respondents below 15 did not receive some of the important questions directed towards adults, these respondents were dropped from this study.

**** Coloured is a term used in South Africa, including on the national census, for persons of mixed race ancestry (CIA – The World Factbook [a], 2015).

***** The respondents were given several more choices, but in this study these have been merged together into four groups of education, using the official division of the education system according to: http://southafrica.usembassy.gov/root/pdfs/study_sa_profile_rev100630.pdf This was partly done in order to simplify the interpretation for readers who do not possess any previous knowledge of the South African education system.

Appendix B- Additional figures and graphs

Figure B.1

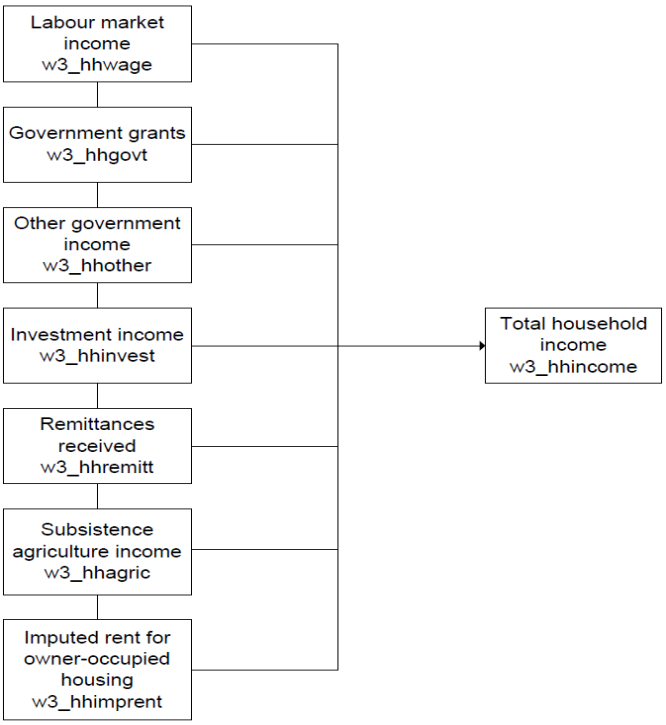

Figure B.1 - *Description of NIDS total Household income, taken from:* *http://www.nids.uct.ac.za/images/documents/wave3/NIDS-W3UserGuide-PublicRelease20150330_v1.5.pdf*

Figure B.2

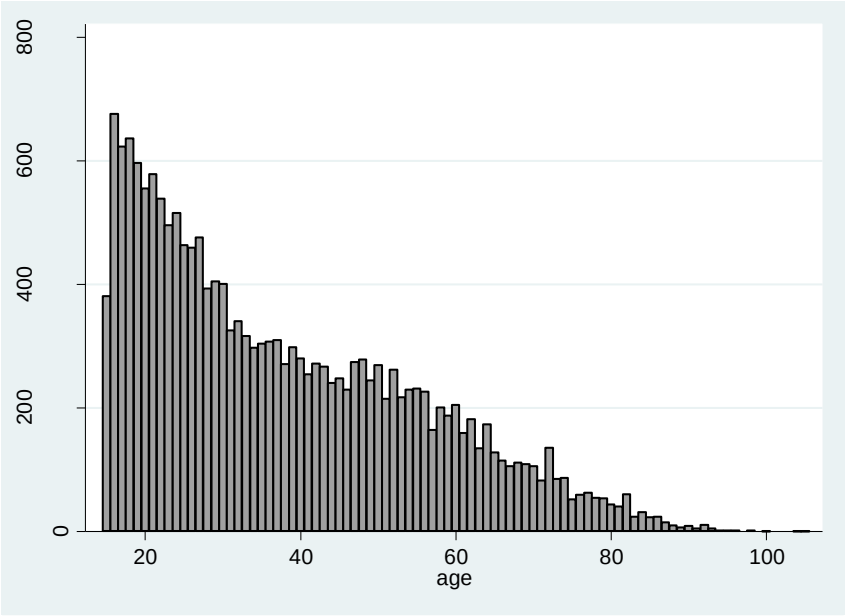

Figure B.2 – *Age distribution within the dataset. Age is displayed on the horizontal axis and the respective frequency of each age is displayed on the vertical axis.*

Figure B.3

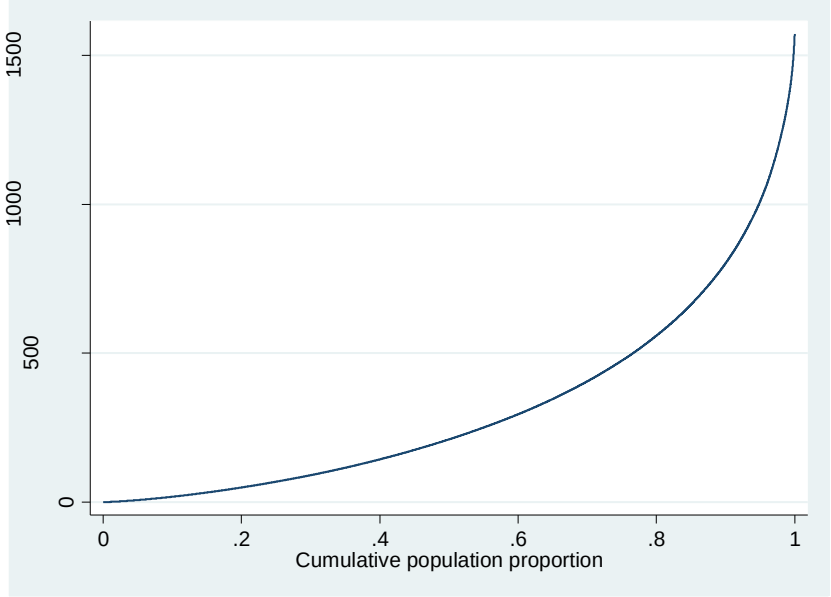

Figure B.3 – *Lorenz curve showing the income distribution in the dataset*