LUND UNIVERSITY
School of Economics and Management

Income, Relative Income and Income Inequality in Australia - What is the Effect on Individual Health?

Josefine Redig

Supervisor
Ulf-G Gerdtham

Master's Thesis
Spring 2014

Abstract

This thesis evaluates the evidence for the absolute-income hypothesis (AIH), relative-income hypothesis (RIH) and income-inequality hypothesis (IIH) in Australia using individual level panel data from the first 11 waves of The Household, Income and Labour Dynamics in Australia (HILDA) Survey under General Release 11. The analysis is performed with an individual fixed effects estimation and the internal validity of the results is addressed in a sensitivity analysis. The results do not support the RIH or IIH. Instead, an adverse effect for the RIH is detected, suggesting a positive effect of relative income on health. Support for the AIH is found at first but the findings are not robust towards the sensitivity analysis, suggesting that the relationship between income and health is endogenous. Based on these findings it could be suggested that a countries institutional setting (for example health care system) has a mediating effect on the relationship between income inequality and health. This needs to be further evaluated in a cross-country study in order to establish a true effect.

Keywords: HILDA, absolute-income hypothesis, relative-income hypothesis, income-inequality hypothesis, health, fixed effects

This paper uses unit record data from the Household, Income and Labour Dynamics in Australia (HILDA) Survey. The HILDA Project was initiated and is funded by the Australian Government Department of Families, Housing, Community Services and Indigenous Affairs (FaHCSIA) and is managed by the Melbourne Institute of Applied Economic and Social Research (Melbourne Institute). The findings and views reported in this paper, however, are those of the author and should not be attributed to either FaHCSIA or the Melbourne Institute.

Table of Contents

1. Introduction

Public health is often seen as one of the main components of overall welfare in most countries. Therefore it is of interest to explore how the surrounding environment affects individual health. In specific, the relationship between income inequality and health has been explored in a vast amount of literature on the topic. At first, a large amount of research found that the relationship between income inequality and health was negatively correlated. Individuals living in countries, or areas, with great differences in income had worse health and were subject to higher mortality rates than individuals living in areas with lower income inequality (Wilkinson, 1996). This notion has also been supported by studies suggesting that income inequality was associated with lower levels of social capital and higher crime rates. This apparent negative effect of income inequality on health is recognized as the income-inequality hypothesis (IIH). However, it has been noted in several studies thereafter that the apparent negative relationship between income inequality and health may be due to a statistical artefact caused by the specific concave relationship between income and health (see for example Rodgers, 1979; Gravelle, 1998 and Wagstaff and van Doorslaer, 2000). Evidence for a concave relationship between income and health, as well as income and survival, has been found in a large amount of research (Wagstaff and van Doorslaer, 2000) and is often called the absolute-income hypothesis (AIH). It has not only been argued that income inequality and individual income has an effect on health, but also that individual *relative* income has an effect on health. This hypothesis is coined the relative-income hypothesis (RIH).

Despite the large amount of literature and research on the topic, evidence supporting the different hypotheses relating income and income inequality to health are mixed. This is a result of that many studies have only utilized aggregate-level data, failing to be able to distinguish between the different hypotheses. In order to do so, individual-level data has to be applied (Wagstaff and van Doorslaer, 2000). As individual-level data has been more frequently applied, evidence supporting the apparent negative effect of income inequality has decreased (Bergh et al., 2013). It is vital to use individual-level data since the different hypotheses have different policy implications (Miller and Paxson, 2006). If it is concluded that it is absolute income rather than income inequality that has an effect on health, policies should be aimed at increasing income for

the poorest individuals in society. However, if it is income inequality that has a detrimental effect on health, policy programmes aimed at decreasing disparities in income can increase health and welfare. If instead it were found that relative income instead of absolute income matters for health, doubling all individuals' income would not make a difference (Miller and Paxson, 2006).

The aim of this thesis is to evaluate the evidence supporting the absolute-income hypothesis (AIH), the relative-income hypothesis (RIH) and the income-inequality hypothesis (IIH) in Australia. This will be done using individual level panel data from the first 11 waves of The Household, Income and Labour Dynamics in Australia (HILDA) Survey under General Release 11. A fixed effects model will be used in the analysis and a sensitivity analysis will be conducted to address the robustness and internal validity of the results.

The thesis will be organized as follows. Section 2 will present the theoretical mechanisms relating income, relative income and income inequality to health. This will be followed by Section 3, which briefly describes previous research on the subject, both in general and more specifically for Australia. Section 4 then describes methodological challenges, advantages, and disadvantages and presents the chosen model. A description of the data used is presented in Section 5. The empirical results are then presented and discussed in Section 6; Section 7 then summarizes and concludes.

2. Theoretical Mechanisms

2.1 Absolute-Income Hypothesis

The effect of income and income inequality on different proxies for health has been explained by several different hypotheses. The definitions of these different hypotheses will in this thesis follow the definitions presented by Wagstaff and van Doorslaer (2000).

The first hypothesis concerns the effect of income on health and is referred to as the AIH. The AIH states that it is income, rather than the direct effect of income inequality, that effects health. It further states that the marginal effect of income on health is diminishing, i.e. the relationship is concave. The implication of this is that average health in a society will increase as average income increases and income inequality decreases (Wagstaff and van Doorslaer, 2000). The notion that income and health are hypothesized to be positively correlated is not controversial. Income enables individuals to provide themselves with food, housing, medical care and physical exercise. Material living conditions and social status that higher incomes can buy increase as income increases (Lundberg et al., 2010). This is also consistent with the predictions from the Grossman model. In the Grossman model, health is both a consumption and production good that yields satisfaction and utility to the individual. As a consumption good, increased income increases the demand for health and health-enhancing goods. At the same time, as a production good, health indirectly yields satisfaction through increased productivity and higher wages (Grossman, 1972). Even though the correlation between income and health is not questioned, it is most probable that the relationship is characterized by reverse causation. Individuals that are healthy can participate in the labour market and thereby health affects income (Lundberg et al., 2010; Bergh et al., 2012). Given that the relationship between income and health suffers from reverse causality, not addressing this problem could yield overestimated effects of income on health (Karlsson et al., 2010).

2.2 Relative-Income Hypothesis

The RIH states that it is the individual's income *relative* to average income of the reference group that has an effect on health. It states that an individual with below-average income will have

worse health (Wagstaff and van Doorslaer, 2000) due to the psychosocial stress associated with having low relative income (Wilkinson, 1996). Differences in income make relatively poor individuals feel stressed and shameful and this in turn reduces individual's general health status.

Even though the RIH in general states that rising average income of an individuals reference group has a detrimental effect on health when keeping individual income constant, explanations to why this may not be the case have also been presented in the literature. Senik (2004) presents an explanation originally derived from Hirschman's tunnel effect presented in 1973. The hypothesis presented by Hirschman (1973) is that if individuals instead of feeling stressed and shameful when comparing their income to the average income of a reference group use reference group income to create positive expectations, the projection of the future is positive and this could then have a positive effect on health.

2.3 Income-Inequality Hypothesis

The third and final hypothesis is the IIH, which states that income inequality in a society directly affects all individuals' health negatively (Wagstaff and van Doorslaer, 2000). This was the original hypothesis presented by Wilkinson (1996) and is the main argument in Wilkinson and Pickett's book *The Spirit Level: Why More Equal Societies Almost Always Do Better.*

Several possible mechanisms of why income inequality could have a detrimental health effect have been presented. The first explanation relating income inequality and its affect on health is related to trust and social capital. Kawachi et al. (1997) find that in societies with larger income inequalities, people are to a greater extent mistrustful. The lack of trust in society due to income inequalities can then erode social capital and lead to a lower degree of social interaction. This can lead to a decrease in health and wellbeing as it has been shown that social interaction can create informal support in cases of crisis and stressful situations. One explanation to why income inequality erodes social trust could be that individuals are most trusting towards individuals that are similar to them. In a society with greater income inequalities, it can be hypothesized that individuals far apart from each other regarding income are not "seen" as similar. Therefore individuals in such societies may be mistrustful and this can be detrimental to the social glue in society that facilitates social interaction (Karlsson et al., 2010).

A second explanation is related to the political mechanisms in a society. Kawachi et al (1997) argue that individuals with higher income pay more (in form of taxes etc.) to the government, than they receive back through for example services and transfers. If this in turn implies that those with higher income have more political power, they could to a larger extent support policies that favour less public spending. Less public spending could imply less provision of public goods and services that could have a potential impact on public health (Krugman, 1996). The hypothesis is thus that income inequality leads to a more polarized society in which fewer common resources are produced (*ibid*). This has been supported by the findings of Kaplan et al. (1996), who find a negative correlation between income inequality and investments in public health in US states. The last explanation hypothesises that societies subject to large inequalities in income are also subject to higher crime and accident rates. This can be a situation that follows due to the first explanation, lack of trust and social capital (Zhao, 2006).

3. Previous Research

3.1 The Ecological Fallacy

There is a large amount of research studying the relationship between income inequality and health. The studies differ in several manners as some use objective measures of health and others use subjective measures of health. However, the most important distinction to make between different studies is regarding whether or not individual-level data has been used. Several studies analysing aggregate data have reached the conclusion that there is strong support for the IIH, finding that income inequality has a detrimental effect on health (see Bergh et al., 2012 for a review). For example, Lynch and others (1998) analyse US data and reach the conclusion that the magnitude of the yearly loss of lives due to income inequality is comparable to the combined magnitude of loss of lives due to diabetes, HIV, lung cancer, motor vehicle accidents, suicide and homicide in 1995. However, with time, researchers have pointed out that the relationship on the aggregate level between health and income may be the result of a concave relationship on the individual level between absolute income and health. This was originally proposed by Rodgers (1979) but is most commonly recognized as Gravelle's "Ecological Fallacy" (1998). Gravelle argued that if there is a concave relationship between health and income on the individual level, individual-level income has to be controlled for. If not controlled for average health will decrease if health losses, for an individual having low absolute income in an unequal society, is greater than the gain in health for individuals with greater levels of income. The concave relationship on the individual level is depicted in Figure 1.

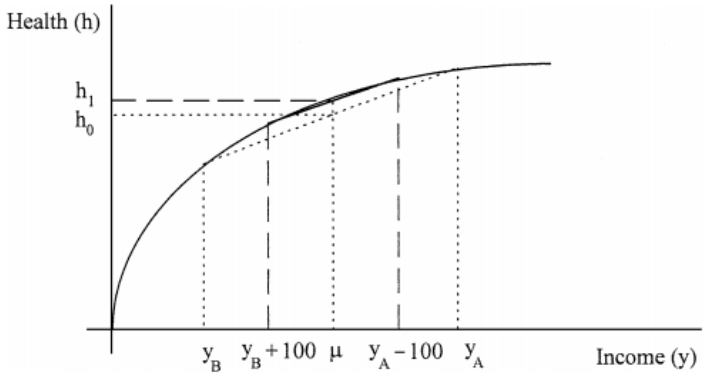

Figure 1 The concave relationship between income and health

Source: Wagstaff and van Doorslaer (2000), Figure 2, p. 546.

To conclude, this implies that a correlation on the aggregate level between income inequality and health does not reveal if it is absolute income, relative income or income inequality that affects health. In order to distinguish between these three hypotheses, which in turn have very different policy implications, individual-level data has to be applied (Wagstaff and van Doorslaer, 2000).

3.2 Individual-level Studies

As more recent research has incorporated individual-level data into the analysis, the evidence regarding the effect of absolute income, relative income, and income inequality on health are not pronounced. Although most studies find support for the AIH when using individual-level data, a majority of the studies do not control for the possible bias due to endogeneity.

Kennedy et al. (1998) and Lochner et al. (2001) examine the impact of income inequality in the US when controlling for individual-level characteristics. Both studies find support for the IIH and further find that those with below-medium and low income experience more detrimental health effects of income inequality. Subramanian and Kawachi (2006) examine the relationship between state-level income inequality and self-rated health in the US. Their results primarily suggest an overall negative impact of income-inequality and they reach the conclusion that it cannot be explained by individual socioeconomic factors. However, their results to a small extent suggest that the detrimental effect on health is greater for advantaged groups. The authors state that this last conclusion is not convincing and that more research is needed in order to clarify the causal effect. Further, Subramanian et al. (2001) and Kahn et al. (2000) find that wealthy individuals in the US have better health when living in states with higher income inequality. Miller and Paxson (2006) analyse the impact of the RIH in the US using semi-aggregated data. By using individual-level income and mortality rates categorized by age, race, gender and place of residence they find a positive effect of relative income and thus reject the RIH. In specific, they find that black men in working ages have lower mortality if living in areas with relatively wealthy neighbours (Miller and Paxson, 2006). Mellor and Milyo (2002) use individual-level data to explore the effect of the IIH in the US and after controlling for geographical and individual-level characteristics, their results do not provide evidence supporting the IIH.

The studies performed in the US to a large extent rely on cross-sectional data. Jones and Wildman (2008) apply British panel data using different functional forms and find that the

chosen functional form is important. They test both the AIH and relative deprivation hypothesis by using a national reference group. Their results from the OLS model supports the AIH and the relative deprivation hypothesis. However, when utilizing a longitudinal approach and controlling for individual unobserved heterogeneity in a fixed effects model, the relationship between health and relative deprivation is no longer significant. The relative deprivation hypothesis is however significant when the effect on psychological well-being for men is analysed in a semiparametric model (Jones and Wildman, 2008). Lorgelly and Lindley (2008) also use British longitudinal data and analyse the evidence for the AIH, IIH and RIH. Their results provide evidence supporting the AIH but not the IIH or RIH when exploring the relationship at county, regional and national level (Lorgelly and Lindley, 2008). Gravelle and Sutton further addressed the results found in Britain in their study from 2009. They conclude that the results are sensitive towards measures used and explore four variants of the RIH in their study. Different measures of income inequality are also addressed in the study and it is concluded that no support for the IIH can be found. The four different variants of the measures for the RIH produce somewhat different results and the authors conclude that overall, only weak evidence supporting the RIH can be found. The negative effect associated with the probability of reporting good health is smaller in magnitude in comparison to the positive effects of income, education, social class and region (Gravelle and Sutton, 2009).

Gerdtham and Johannesson (2004) use Swedish register data and they find that mortality decreases as individual income increases, but reject a negative association between income inequality and mortality as well as relative income and mortality. Income inequality and relative income is measured on the municipality level, assuming then that individuals reference group are neighbours. One explanation to the results has been that the Swedish welfare state can compensate for the negative effects that otherwise would be present, a similar explanation to the results found in the UK was put forward by Lorgelly and Lindley (2008). Dahl et al. (2006) use Norwegian data to examine the effect of income inequality on mortality. In contrast to Gerdtham and Johannesson (2004), they find in their multi-level analysis that income inequality has a positive effect on mortality. The conclusion that there is no income inequality effect in egalitarian countries and Nordic welfare states is thus not supported by Dahl et al.'s study.

3.3 The Australian Context

The amount of studies analysing the different hypotheses within Australia are few. There are however several studies with the aim of evaluating evidence supporting the AIH. The results from these studies indicate that there is evidence supporting the AIH in Australia. Cai (2009) finds support for the AIH when addressing the problem of endogeneity between income and health for the elderly. Similar results are found by Chotikapanich et al. (2003), who conclude that ill-health is concentrated among lower income groups.

Regarding studies analysing relative income and income inequality in Australia, the results presented in this thesis will mainly be compared with the paper by Bechtel el al. (2012). They utilize the first eight waves of the HILDA survey and analyse the relationship between mental health and income inequality in Australia, using a fixed effects model, by applying different measures of income inequality on the neighbourhood level. Bechtel el al. further tests the relative deprivation hypothesis, which "detects an adverse association for some individuals within a group" (Bechtel el al., p. 5, 2012). This is tested by including the average of the income of the difference in incomes of all individuals who have a higher income. The chosen reference group for this measure is the neighbourhood. Their results do not support the IIH but they do find support for the relative deprivation hypothesis, although the effect seems to be small in magnitude. Further, they do find support for the AIH but conclude in their sensitivity analysis that the relationship between income and health is not robust, suggesting that the relationship is partially driven by reverse causality (Bechtel el al., 2012).

To conclude, results from previous studies differ. Different methods have been used to evaluate effects, different definitions of hypotheses have been used and different levels of analysis have been applied. However, when examining the potential effects present, the specific context in which the question is analysed should be considered. Countries' or regions' social and economical framework most probably has an impact on the potential effects. It has been suggested that evidence supporting the IIH is often found in the US due to its institutional settings. Likewise, it has been hypothesized that the absence of evidence supporting the IIH in countries like Sweden and the UK is due to their egalitarian properties as welfare states. For example, Grönqvist et al. (2012) suggest that the services produced by and the structure of the welfare state in Sweden compensates for the potential detrimental health effect of income inequality. Bechtel el al. (2012) also suggests that health services produced by the state may mediate the effect of income

inequality. Health care in Australia is universal and provided by both public and private providers. In 2011, health care expenditure was equal to 9,2% of GDP, whereas 68,4% was publically financed. The corresponding amounts as % of GDP in 2011 for Sweden, UK and the US were 9,5%, 9,4% and 17.7%. The amounts publically financed in these countries were 81,6% for Sweden, 82,8% for UK and 47,8% for the US (World Bank, 2014).

4. Methodology

4.1 Income, Relative Income and Income Inequality

In order to test the AIH, the log of income will be used to allow for a non-linear relationship between income and health. This is the standard specification used in similar studies (see for example Bechtel el al., 2012).

To measure income inequality, the Gini coefficient will be used. The Gini coefficient is a common measure of income inequality and is defined as "twice the area between the Lorenz curve and the egalitarian line of perfect equality". The Gini coefficient has a value between zero and one, where zero represents perfect equality and one perfect inequality. The calculations of the Gini coefficients have been calculated using equivalised household disposable income since this is the definition that represents "true" income inequality (a more detailed description of the calculation of this income measure can be found in Section 5.2). Gini coefficients will be calculated for the different major statistical regions provided in the data sample.[1] In total, there are 13 major statistical regions in Australia: Sydney, Balance of New South Wales, Melbourne, Balance of Victoria, Brisbane, Balance of Queensland, Adelaide, Balance of South Australia, Perth, Balance of Western Australia, Tasmania, Northern Territory and Australian Capital Territory. This implies that income inequality in the major cities (Sydney, Melbourne, Brisbane, Adelaide and Perth) has one value whereas for those individuals not residing in these major cities have a separate value calculated based on the income of other individuals in the region also not residing in these cities.

Including the mean income of the individuals reference group will test the RIH. The reference group refers to the group of individuals that an individual compares him or herself with. Choice of reference group can be based on social, demographic or geographical grounds. Previous studies have to a large extent used individuals in the same age group (for example Karlsson, 2010) or neighbourhood (for example Bechtel el al., 2012; Gerdtham & Johannesson, 2004) as a reference. In this thesis, the reference-group considered are colleagues employed in the same

[1] Gini coefficients have been calculated using Stata's *inequal* command.

industry according to the 3.2 2-digit International Standard Industry Classification (ISIC). Support for using colleagues or co-workers as a reference group can be found in a paper by Clark and Senik (2010), who with survey information from 18 European countries analyse "Who compares to whom?". The results from their paper suggest that colleagues are the most frequently stated reference group (Clark and Senik, 2010). In the data set applied in this thesis, current industry is only available for individuals during years that they have stated that they are employed, individuals who are unemployed or not in the labour force are therefore missing a reference group. For those unemployed, the mean income of other unemployed will be used and for those who are not in the labour force, the mean income of others not in the labour force is used.

When considering the relationship between these different measures and health, there are several potential sources of endogeneity present. The first issue that arises that is a potential threat to the validity of the model is measurement errors in the explanatory variables. If there are measurement errors in the explanatory variables included, the coefficients will be biased towards zero (attenuation bias). Measurement errors arise when something cannot be measured absolutely accurately, due to for example reporting or coding errors, or when calculations are made on a small number of observations. Statistically, measurement errors are not a problem per se as long as coefficients are interpreted as the effects of the *reported* variables. However, measurement errors are problematic if one wants to interpret the underlying value of the coefficients (Verbeek, 2010).

Further, there are several identification problems that arise due to endogeneity. Establishing a causal relationship between income and health is complicated due to reverse causality. If the relationship between income and health is characterized by reverse causality, estimates from an ordinary regression model will be biased, inconsistent and overestimated (Verbeek, 2010; Kennedy, 2008). An additional source of endogeneity bias can arise due to omitted variable bias. Omitted variable bias arises when relevant explanatory variables that are correlated with the other explanatory variables are omitted from the model. Omitted variable bias is also a problem when there are unobservable factors that influence the other explanatory variables. In this context it means that it is not possible to observe or control for individual specific factors that possibly could influence individual's possibilities to produce health. This factor is heterogeneous for all individuals and often difficult to observe and referred to as individual unobserved heterogeneity

(Verbeek, 2010). In an ordinary OLS model, it is not possible to take this into account. Establishing a causal relationship from the included explanatory variables is in such a case not correct as the estimated effect may be correlated, and thus caused by, an unobserved factor (*ibid*).

4.2 Fixed Effects Model

To estimate the effect of income, relative income and income inequality on health, an individual fixed effects model will be applied. It is advantageous to use a fixed effects model when it could be assumed or suspected that the explanatory variables are not exogenous, i.e. when individual unobserved heterogeneity is correlated with one or more of the explanatory variables. This means that we can control for unobserved but fixed omitted variables (Angrist and Pischke, 2009), capturing all unobserved individual variation that is fixed over time (Verbeek, 2010). Without using a fixed effects model, this causes omitted variable bias and a conventional OLS regression will give biased and inconsistent estimates (Verbeek, 2010).

At first, an OLS regression will be estimated including only the main variables of interest; income, relative income and income inequality (Model 1). Thereafter, the same model will be estimated with the fixed effects estimator (Model 2) and control variables will be added in Model 3. In order to examine the differential effects for men and women, the fixed effects estimator will be estimated for sub-samples of each sex (Model 4 and 5).

Even though there are advantages of applying an individual fixed effects model, there are several limitations and disadvantages with the model. The fixed effects estimator is often called the within-estimator since the model only uses within-individual variation. Therefore, since the variation is equal to zero for all time-invariant characteristics, all time-invariant characteristics (both observed and unobserved) are eliminated. A result of this is that it is impossible to estimate the effect of such time-invariant characteristics, such as country of birth, gender and other variables that do not change during the observed time period (Andreß et al., 2013). An additional assumption of the fixed effects model is that the regressors need to be strictly exogenous. This assumption holds if the unobserved factors do not vary over time, but if they do, the strict exogeneity assumption is violated (Angrist and Pischke, 2009). Further, the model assumes that individual unobserved heterogeneity is correlated with the explanatory variables. If the individual component, i.e. the individual unobserved heterogeneity, does not correlate with other

explanatory variables, there is a loss of efficiency and a random effects model is favourable. An additional pitfall of the fixed effects approach is that the attenuation bias due to measurement errors is greater than in OLS models (Andreß et al., 2013).

4.3 Endogeneity

Even though the fixed effects model can deal with some of the potential problems with endogeneity, there are some sources of endogeneity that the model is not robust towards and this will be addressed in a sensitivity analysis (Models 6-11). It should be noted that the sensitivity analysis is conducted with the fixed effects estimator.

The chosen reference group considered for those who are unemployed or not in the labour force is an assumption that can be discussed whether or not it is valid. The model will therefore be reestimated including only individuals who are employed, i.e. individuals who have a reference group categorized by industry (Model 7). As mentioned in Section 4.1, the fixed effects estimator is particularly sensitive towards measurement errors. Therefore, the model will be estimated excluding reference groups with fewer than 100 observations in order to address the robustness of the estimated effect of the RIH (Model 8). In this specification, individuals who are unemployed or not in the labour force will be excluded as well since the assumption regarding their reference groups may be prone to measurement error, as argued above. Gerdtham and Johannesson (2004) apply such a method, i.e. exclude regions with few observations, in their sensitivity analysis towards their measure of the Gini coefficient. However, the Gini coefficient in this thesis is measured for a larger geographical area and the number of observations used when calculating Gini coefficients is fairly large and therefore there is no reason to do this for the Gini coefficient.

As it has previously been argued, the direction of causation between income and health runs from both income to health and vice versa. Therefore, an attempt to address the robustness of the results regarding income and health will be made. Bechtel et al. (2012) argue that using household income rather than individual income reduces the bias in the estimates that can arise due to reverse causality. The reason for this is that even though an individual is sick or unable to work, other members of the household can still earn money. However, this is not adequate enough to eliminate problems with endogeneity and therefore the same method of sub-sampling

used in Bechtel et al. (2012) will be applied here as well. The aim of the sub-sampling is to remove some of the effect that health can have on income. The first sub-sample will exclude individuals that have been unemployed in any of the 11 waves (Model 9). The second sub-sample will estimate the model by first excluding all individuals who have taken unpaid sick leave during the past 12 months (Model 10), and thereafter estimate the model excluding only all individuals who have had more than 5 unpaid days of sick leave during the past 12 months (Model 11). This will eliminate individuals who have had low income because of not being able to work when sick.

5. Data

5.1 The HILDA Survey

The HILDA survey is a longitudinal household-based survey that started in Australia in 2001, which collects information concerning economic and subjective wellbeing, labour market dynamics and family dynamics. All members in a household are included in the scope of the data but only those aged 15 and older are interviewed. The survey population included in the study is representative of the Australian population with the exception of individuals situated in remote areas of the country. The initial wave consisted of 7 682 households and in total 19 914 individuals (Summerfield et al., 2012). Due to attrition, not all the initial 19 914 interviewed individuals are included in the data set analysed in this thesis. The constructed data set used for this thesis is a balanced panel data set using data from waves 1-11 under General Release 11. The sample covers in total 7 229 interviewed individuals aged 15 years and older, resulting in 79 519 observations. However, as there are some observations with missing values on reported health, the data set is essentially unbalanced. Missing values are due to nonresponse, no self-completed questionnaire, multiple responses or not being able to determine the given response. The unbalanced panel was chosen for the main model since the fixed effects model relies on within-variation, and if this within-variation is small, a large sample is needed to prove any significance of the explanatory variables (Andreß et al., 2013). In order to address if the unbalanced panel estimates were biased in comparison to using a balanced sub-panel, a balanced sub-panel where only individuals with complete health data for all 11 waves (Model 6) is included in the sensitivity analysis and will be further discussed in Section 6.2.

5.2 Income and Health Measures

In order to measure individual health, the SF-6D health state classification provided under the General Release dataset is used. This variable is derived from the SF-36 Health Survey instrument that is included in the questionnaire. The SF-36 consists of 36 different items across eight different health components: physical functioning, role-physical, bodily pain, general health, vitality, social functioning, role-emotional and mental health (Summerfield et al., 2012). These eight different components are transformed into an index between 0-100. The SF-6D health state

classification is thereafter derived by applying preference-based measures using utility weights (see Brazier et al., 2002), in order to apply different utilities to different health states. In total there are 18 000 different health states. The result is a measure of self-assessed health along a scale from 0 to 1, where 1 is equal to full health.

Individual income is calculated using household financial year disposable regular income. This measure is a sum across all household members gross regular income (wages and salaries, investment, business, private pension, and Australian government transfers) minus government taxes (Summerfield et al., 2012). The measure of income is transformed according to the Organisation of Economic Co-operation and Development modified equivalence scale in order to allow comparisons between households of different size and over time. This is a common transformation and it is calculated by devoting additional adults in a household a weight equal to 0.5 and each child under the age of 15 a weight of 0.3. This implies that the income for single-person households is not modified. Even though only households that have been interviewed in each wave are kept in the data set analysed in this thesis, missing income data still appears. However, the HILDA survey team uses various different methods to impute missing income data and therefore household financial year disposable regular income is not missing for any observation in the data sample (Summerfield et al., 2012).

5.3 Control Variables

In addition to the measures of health, income and income inequality, several control variables are included in the model. The number of dependents an adult has, aged 0-14 years, is included. A dummy variable is incorporated to control for if the individual is employed in the labour force or not. The dummy variable is equal to 1 if the individual is employed and 0 in other cases (if the individual is unemployed or not in the labour force). If the individual is married or in a de facto relationship they receive the value 1 on an additional dummy variable controlling for marriage/ cohabiting status. Individuals who are divorced, separated, widowed or never married and not de facto are given a 0 for this variable. A dummy variable indicating whether or not the individual has completed higher education is included. The dummy variable is equal to 1 if the individual has a bachelor, diploma, honours or doctorate degree. In other cases it is equal to 0.

5.4 Sensitivity Analysis

The sensitivity analysis presented in Section 4.3 will utilize variables related to characteristics of the individual's employer and the amount of unpaid leave that the individual has had during the past 12 months. Specifically, in order to determine if the individual has had unpaid sick leave, three variables from the HILDA survey are utilized. If the individual has answered yes to the question "Have you taken any unpaid leave in the last 12 months?" and no to the question "Does your employer provide paid sick leave?", they are excluded from the sample in Model 10. Further, the days of unpaid leave in the last 12 months is used to determine if the individual has had more than 5 days of unpaid leave in the last 12 months (Model 11). As these three questions only were asked form wave 5 and onwards, Model 10 and 11 only use data from waves 5-11 in the estimation.

6. Results

6.1 Empirical Results

The results from the main estimations can be found in Table 1 (Model 1-5) and results from the sensitivity analysis are presented in Table 2 (Model 6-11). All estimations are conducted in the statistical software package Stata, version 12, and all models are estimated with robust standard errors.[2] Descriptive statistics for the main variables of interest can be found in Appendix.

Table 1 Main Results- all except Model 1 estimated with the fixed effects estimator

Variable	*Model 1* OLS without control variables *N=70 603*	*Model 2* Without control variables *N= 70 603*	*Model 3* With control variables *N= 69 316*	*Model 4* Sub-sample: men *N=31 674*	*Model 5* Sub-sample: women *N=37 642*
Constant	0.4816*** (0.0100)	0.7963*** (0.0105)	0.7982*** (0.0126)	0.8016*** (0.0182)	0.7980*** (0.0175)
Gini	-0.2221*** (0.0194)	-0.1272*** (0.0231)	0.0302 (0.0239)	0.0441 (0.0342)	0.0157 (0.0333)
Mean income_ref	$1.22x10^{-6}$*** ($4.22x10^{-8}$)	$9.53x10^{-8}$** ($4.55x10^{-8}$)	$3.87x10^{-7}$*** ($8.37x10^{-8}$)	$2.84x10^{-7}$*** ($1.10x10^{-7}$)	$5.15x10^{-7}$*** ($1.29x10^{-7}$)
log(income)	0.0293*** (0.0009)	0.0001 (0.0009)	0.0022** (0.0009)	0.0018 (0.0013)	0.0026** (0.0012)
Age			-0.0021*** (0.0002)	-0.0019*** (0.0003)	-0.0023*** (0.0003)
Employed			0.0105*** (0.0020)	0.0136*** (0.0032)	0.0079*** (0.0027)
Education			0.0060* (0.0037)	-0.0012 (0.0056)	0.0123** (0.0048)
Married/cohabiting			0.0087*** (0.0018)	0.0140*** (0.0134)	0.0053** (0.0024)
Dependents			0.0040*** (0.0008)	0.0009 (0.0011)	0.0065*** (0.0010)

Significance level: ***p-value<0.01, ** p-value<0.05, * p-value<0.1. Robust standard errors in parentheses.

[2] Clustered-robust standards errors are recommended when using panel data (Verbeek, 2010), with fixed effects this is identical to using robust standard errors.

The results from the initial OLS model, including only the Gini coefficient, mean income of the reference group and individual income estimates a negative and significant impact between individual health and the Gini coefficient equal to -0.2221, implying that income inequality exerts a detrimental effect on health. Mean income of the reference group exerts a positive and highly significant effect equal to $1.22x10^{-6}$ on individual health. Further, the coefficient for individual income is significant at a level of 1% and equal to 0.0293. In the second model, when the fixed effects estimator is applied, the effects all decrease in magnitude. The effect of the Gini coefficient is still negative and significant at a level of 1%, but has decreased to -0.1272. The initially small effect of mean income in the reference group decreases further, to $9.53x10^{-8}$ but remains significant at a level of 5%. The fixed effects estimate for individual income is also positive but insignificant. Model 3 is then extended by adding control variables to the specified model. When this is done, the estimated effect of the Gini coefficient differs. The estimated coefficient is now equal to 0.0302 and is insignificant, implying that no statistically significant relationship between the Gini Coefficient and health can be detected in the model. The estimated coefficient of individual income increases slightly from 0.0001 to 0.0022 and is now significant at a level of 5%, implying that increased income enhances health. The estimated coefficient for mean income of the reference group increases in magnitude and significance, from $9.53x10^{-8}$ at a significance level of 5% to $3.87x10^{-7}$ at a level of 1%. Higher mean income in the reference group is thus associated with better health.

Concerning the control variables added in Model 3, individuals' age has a negative and significant impact on health. An additional year of ageing decreases health with 0.0021 units. The estimated coefficient for the dummy variable indicating whether or not the individual is married or cohabiting is positive and significant. The effect on health due to marriage or cohabitating is equal to 0.0087. The coefficient for the dummy variable indicating if the individual is employed or not indicates that health significantly increases with 0.0105 units when an individual is employed. The estimated effect of having high education is positive and significant. When the individual has higher education, health increases with 0.006 units. The last control variable added to the specification is the number of dependents the individual has. An additional dependent increases an individuals' health with 0.004 units. To conclude, these results imply that being employed, married or cohabiting, having high education and dependents enhances health. Ageing however has a detrimental effect on health. These results are consistent with expected findings.

Models 4 and 5 have the same specification as Model 3 but are run on sub-samples of each sex. The conclusion of the estimated results regarding the Gini coefficient are in both of these models the same as in Model 3; the Gini coefficient is estimated to have a positive but insignificant effect on health. The estimated effects regarding mean income of the reference group differ slightly from Model 3 and the effect seems to be much larger for women than men. The estimated effect of individual income is positive and significant for women but insignificant for men and also here the effect appears to be larger for women than men. Regarding the covariates, the main results found for age, employed, and married or cohabiting apply in Model 4 and 5 as well but are slightly different in magnitude. For the sub-sample estimated for women, results are similar as in Model 3 for the coefficients of education and number of dependents as well. For men however, there is no statistically significant relationship found between the number of dependents or education and health.

Table 2 Fixed Effects Model- Sensitivity Analysis

Variable	*Model 6* Balanced sub-sample *N=39 008*	*Model 7* Excluding unemployed or not in the labour force *N=45 109*	*Model 8* Excluding reference groups with less than 100 observations *N=44 292*	*Model 9* Excluding individuals unemployed in *any* wave *N=59 823*	*Model 10* Excluding individuals with unpaid sick leave (only waves 5-11) *N=37 065*	*Model 11* Excluding individuals with more than 5 unpaid sick days (only waves 5-11) *N=39 773*
Constant	0.7879*** (0.0164)	0.7800*** (0.0146)	0.7853*** (0.0148)	0.8060*** (0.0139)	0.8586*** (0.0194)	0.8512*** (0.0185)
Gini	0.0149 (0.031)	0.0247 (0.0291)	0.0196 (0.0292)	0.0352 (0.0255)	0.0349 (0.0332)	0.0284 (0.0318)
Mean income_ref	$2.27x10^{-7}$** ($1.08x10^{-7}$)	$2.34x10^{-7}$** ($9.42x10^{-8}$)	$2.78x10^{-7}$*** ($9.9x10^{-8}$)	$3.97x10^{-7}$*** ($9.15x10^{-8}$)	$3.67x10^{-7}$*** ($1.08x10^{-7}$)	$3.82x10^{-7}$*** ($1.03x10^{-7}$)
log(income)	0.0025** (0.0011)	0.0031*** (0.0012)	0.0029** (0.0012)	0.0019** (0.0010)	0.0010 (0.0012)	0.0013 (0.0011)
Age	-0.0015*** (0.0003)	-0.0013*** (0.0003)	0.0014*** (0.0003)	-0.0021*** (0.0002)	-0.0027*** (0.0102)	-0.0027*** (0.0003)
Employed	0.0092*** (0.0026)			0.0102*** (0.0024)	0.0083*** (0.0031)	0.0073** (0.0029)
Education	0.0054 (0.0046)	0.0048 (0.0040)	0.0049 (0.0041)	0.0778* (0.0041)	-0.0108 (0.0114)	-0.0061 (0.0066)
Married/cohabiting	0.0080*** (0.0026)	0.0089*** (0.0021)	0.0090*** (0.0021)	0.0010*** (0.0020)	0.0053* (0.0029)	0.0048* (0.0027)
Dependents	0.0037*** (0.0010)	0.0036*** (0.0008)	0.0036*** (0.0008)	0.0041*** (0.0008)	0.0026** (0.0013)	0.0027** (0.0012)

Significance level: *** p-value<0.01, ** p-value<0.05, * p-value<0.1. Robust standard errors in parentheses.

The results from the sensitivity analysis are displayed in Table 2. First, the model was reestimated using a balanced sub-sample of the panel. When comparing these results with the results from Model 3, no major differences can be detected. Coefficients for the main variables of interest are in large the same as in Model 3. The coefficient for individual income has increased from 0.0022 to 0.0025 and the coefficient for mean income in the reference group has decreased somewhat from $3.87x10^{-7}$ to $2.27x10^{-7}$. Regarding the Gini coefficient, the coefficient is not either here significant but has decreased to 0.0149. The coefficients for the control variables are marginally

different in magnitude in comparison to Model 3 but the conclusions regarding the controls are the same except for education, since the coefficient no longer is significant at a level of 10%.

The specification in Model 7 uses the unbalanced panel used in Models 2-3 but excludes individuals when they are unemployed or not in the labour force to address the robustness of the variable for mean income of the reference group. The coefficient for mean income in the reference group is in this specification positive and significant. In comparison to Model 3, the coefficient has decreased somewhat in size, from $3.87x10^{-7}$ to $2.34x10^{-7}$ and is significant at a level of 5% instead of 1%. To further account for measurement errors in the variable accounting for mean income in the reference group, the restriction for Model 7 is extended and in addition to excluding individuals who are unemployed or not in the labour force, reference groups with less than 100 observations are excluded in Model 8. The coefficient increases somewhat but is still very small in size, however it increases in significance to a significance level of 1% (as in Model 3).

Models 9-11 are specified in order to address if there is reason to suspect endogeneity between income and health. First, Model 9 excludes individuals who have been unemployed in any wave from the sample. Comparing to Model 3, the coefficient for individual income decreases from 0.0022 to 0.0019. The level of significance remains the same (5%). In Model 10, when individuals with unpaid sick leave are excluded from the sample, the coefficient for individual income is no longer significant. This is true also in the last specification when the restriction of excluding individuals with unpaid sick leave is relaxed and only individuals with more than 5 days of unpaid sick leave are excluded in Model 11.

6.2 Discussion of Results

In this analysis, no robust evidence supporting the AIH, RIH or IIH can be found. At first, when control variables are not included in the model and the model is estimated with OLS, evidence supporting the IIH and AIH is found, as the Gini coefficient seems to exert a significant detrimental effect on health and individual income a significant positive effect. When the model instead is estimated with the fixed effects estimator, taking into account individual unobserved heterogeneity, the magnitudes of the coefficients significantly decrease. The Gini coefficient still

seems to exert a detrimental effect on health in the first fixed effects model, but individual income has lost significance.

When individual-level control variables are added to the model, the significant effect of the Gini coefficient disappears and the coefficient is positive. This highlights the importance of controlling for individual-level characteristics, both observed and unobserved. There does not either appear to be an effect of income inequality on health when the effect is estimated for men and women separately. Individual income has a significant positive effect on health when control variables are included in the regression and this is true when the model is estimated for only women. However, this relationship is not robust for men or when the endogeneity between income and health is addressed. Not in any specification can evidence for the RIH be found, but adverse evidence is found in all specifications, as the coefficient for mean income in the reference group is positive and significant at a level of at least 5%. It should however be noted that the size of the coefficient is basically equal to zero and this should be taken into consideration when assessing the effect of relative income on individual health.

The study by Bechtel el al. (2012) is the existing study that is most comparable to this one, since it utilizes the same data (but fewer waves) and methodology. The results found in this analysis regarding the IIH and AIH comply with the results found in their study. Other studies that find evidence rejecting the IIH are amongst others Meller and Milyo (2002), Lorgelly and Lindley (2008) and Gravelle and Sutton (2009). The adverse effects detected for the RIH are consistent with findings by Gerdtham and Johannesson (2004) as well as Miller and Paxson (2006).

The main estimations in this analysis utilized an unbalanced sample. Therefore, it was in the sensitivity analysis analysed whether or not using a balanced sub-sample would give different estimations. However, it can be concluded that aside from minor changes in size and significance for some of the variables, the conclusions regarding the main variables of interest do not differ and thus the results are not prone to major bias when using the unbalanced sample. It should however be noted that since the data set first constructed was intended to be a balanced sample, the average number of observations per individual is high (9.8 whereas 11 is the maximum). The sample is unbalanced due to missing values for health data, not due to attrition from the survey. If an unbalanced sample had been constructed, including observations from individuals who have dropped out from the survey, this could give different results, as it may be the case that

individuals who are kept in the sample all 11 waves possess certain characteristics. If this is the case, the estimated effects found in this study may suffer from selection bias. However, the fixed effects estimator allows this selection to depend on the individual unobserved heterogeneity component, as long as it is constant over time for each individual. The bias arising from selection bias is thus smaller in the fixed effects estimator than in the random effects estimator (Andreß, 2013).

As already mentioned, the choice of reference group for those who are unemployed or not in the labour force is one assumption that could be subject to criticism. In the empirical analysis it is assumed that when an individual becomes unemployed, he or she will consider other unemployed as a reference group. The same is assumed for individuals not in the labour force. This could be the case, but it could also be the case that the individual continues to consider former colleagues and co-workers as a reference group. To reduce the potential bias that arises from incorrect specification of the reference group for those who are unemployed or not in labour force, they are in one specification excluded from the sample. Although the coefficient decreases in size, the change in effect is negligible due to the initially small size. Further, in an additional specification, reference groups with fewer than 100 observations are also excluded. However, the results regarding the effect of mean income in the reference groups are insensitive towards this and only marginally differ in magnitude.

The sensitivity analysis addressing the problem with reverse causality between income and health confirms that there indeed is reason to suspect endogeneity. One part of Bechtel el al.'s (2012) sensitivity analysis is replicated and extended in this thesis. When individuals with unpaid sick leave are excluded from the sample, the effect of individual income on health is no longer significant. This conclusion remains when only individuals with more than 5 days of unpaid sick leave are excluded. When these individuals are excluded, it is intended to remove some of the individuals that have had an effect on income due to ill health, decreasing the causation from health to income. In both specifications, individual income is no longer significant, implying that the effect of income on health has been overestimated in the other specifications due to reverse causality. These findings confirm the results found by Bechtel el al. (2012).

Regarding the theoretical mechanisms for the different hypotheses presented in Section 2, support for the AIH is found at first in the analysis presented here. This relationship is not

robust in the sensitivity analysis, suggesting that there is an effect of health on income. However, this needs to be further explored to determine the exact relationship, as the only thing the coefficient in this analysis can tell is that there appears to be a correlation.

The RIH cannot be supported in any of the specifications in this analysis. This can depend on many things and may be a result of not correctly specifying the correct reference group. However, if the reference group considered in this analysis is correct, Hirschman's tunnel effect presented in Section 2 could be an explanation to the apparent positive effect of mean income in the reference group. An additional explanation could be that the categorical industry groups with higher mean income are characterized with characteristics that are beneficial for health. This reasoning has been used when positive effects of regional mean income have been found (Miller and Paxson, 2006). Geographical regions with higher mean income can have positive externalities that have positive effects on public health. This effect could be larger than the detrimental effect of feeling ashamed and stressed when comparing income to colleagues and co-workers (the RIH) in the same industry and therefore a negative effect may not be detected, it could of course also be the case that the detrimental effect is not present at all.

Evidence for the IIH is not found in any of the specifications including control variables. This could be a result of the Australian health care system, as suggested also by Bechtel el al. (2012). In absence of the Australian state's services, income inequality could exert an effect on public health. The results found in this study could also be a result of measuring income inequality on a level of aggregation in which there is no such effect. However, Bechtel el al. (2012) tests if income inequality over a smaller geographical region has an effect and their results do not support the IIH. Further, different measures of income inequality are not assessed in this analysis. Bechtel el al. (2012) assesses several different indices of income inequality in their study and concludes that their results are robust towards the type of measure applied. Therefore, it is concluded that this is probably not a major source of bias in this empirical analysis.

7. Conclusion

The aim of this paper was to empirically investigate the evidence for the AIH, the RIH and the IIH in Australia using data from the first 11 waves of the HILDA survey when applying the fixed effects estimator. Further, the robustness of the results was addressed in a sensitivity analysis. Results from the fixed effects estimations did not provide evidence supporting the RIH and IIH. An adverse effect for the RIH was detected, suggesting a positive relationship between relative income and health. Evidence supporting the AIH was found but this was not robust in the sensitivity analysis, implying that there is an endogenous relationship between income and health.

Findings of previous research to a large extent differ and it is clear that the choice of methodology and model specifications give different results, which is also proved in this study. One reason to why specifications may differ could be that there is no consensus regarding the exact relationship between health and income inequality and the hypotheses presented so far consider aspects over a large spectrum considering political mechanisms, social structures as well as psychological phenomena (Bergh et al., 2013). In large, evidence supporting the IIH and various definitions of the RIH are ambiguous. Evidence supporting the AIH is rather large, but most studies fail to account for reverse causality between income and health. Findings in this paper suggest that this leads to bias and overestimated effects of the AIH. By using an appropriate instrument for income, the problem of endogeneity between income and health can be further evaluated in an instrumental variable analysis.

The findings in this study do not find evidence for the IIH in Australia, implying that no evidence is found supporting policies aimed at decreasing income inequality. However, even though no detrimental effects can be found, it should be mentioned that redistribution of income leads to redistribution of other resources and that there are other goals with redistributing income in society (Lundberg et al., 2010; Bechtel el al., 2012). Examples of this are for example the provision of public services such as health- and childcare.

In future research, in order to evaluate if the health care system in a country has a mediating effect on the effect of income inequality on health, cross-country studies need to be conducted to

analyse the impact of countries health care systems and other institutional characteristics. Further, it should be investigated to a larger extent *if* individuals compare themselves to others and to whom in that case to be able to properly assess the RIH. This is a challenge as it probably differs for both different individuals and different countries (Miller and Paxson, 2006; Karlsson et al., 2010).

References

Andreß H.J., Golsch K. and A.W. Schmidt (2013), *Applied Panel Data Analysis for Economic and Social Surveys,* Springer.

Angrist J.D and J.S. Pischke (2009), *Mostly Harmless Econometrics, An Empiricist's Companion,* Princeton University Press, Oxfordshire.

Bechtel L., Lordan G. and D.S Rao (2012), Income inequality and mental health- empirical evidence from Australia, *Journal of Health Economics,* 21(1): 4-17.

Bergh A., Nilsson T. and D. Waldenström (2012), *Blir vi sjuka av inkomstskillnader? : En introduktion till sambanden mellan inkomst, ojömlikhet och hälsa*, Studentlitteratur AB.

Brazier J., Roberts J. and M. Deverill (2002), The estimation of a preference-based measure of health from the SF-36, *Journal of Health Economics,* 21(2): 271–292.

Cai L. (2009), Be wealthy to stay healthy: an analysis of older Australians using the HILDA survey, *Journal of Sociology,* 45(1): 55-70.

Chotikapanich D., Creedy J. and S. Hopkins (2003), Income and Health concentration in Australia, *The Economic Record,* 79(246): 297-305.

Clark A.E. and C. Senik (2010), Who compares to whom? The anatomy of income comparisons in Europe, *The Economic Journal,* 120(544): 573-594.

Dahl E., Elstad I.J., Hofoss D. and M. Martin-Mollard (2006), For whom is income inequality most harmful? A multi-level analysis of income inequality and mortality in Norway, *Social Science and Medicine,* 63(10): 2562-2574.

Gerdtham U. and M. Johannesson (2004), Absolute income, relative income, income inequality, and mortality, *The Journal of Human Resources*, 39(1): 228–247.

Gravelle H. (1998), How much of the relation between population mortality and unequal distribution is a statistical artefact? *British Medical Journal*, 316: 382-285.

Gravelle H. and M. Sutton (2009), Income, relative income and self-reported health in Britain 1979-2000, *Health Economics,* 18(2): 125-145.

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

Grönqvist H., Johansson P. and S. Niknami (2012), Income inequality and health: Lessons from a refugee residential assignment program, *Journal of Health Economics,* 31(4): 617-629.

Hayes C. and N. Watson (2009), *HILDA Imputation Methods*, HILDA Project Technical Paper Series No. 2/09*,* Melbourne Institute of Applied Economic and Social Research.

Hirschman A. (1973), The changing tolerance for income inequality in the course of economic development, *Quarterly Journal of Economics*, 87(4): 544–566.

Jones A.M. and J. Wildman (2008), Health, income and relative deprivation: evidence from the BHPS, *Journal of Health Economics,* 27(2): 308-324.

Lorgelly P.K. and J. Lindley (2008), What is the relationship between income inequality and health? Evidence from the BHPS, *Health Economics,* 17(2): 249-265.

Senik C. (2004), When information dominates comparison: learning from Russian subjective panel data, *Journal of Public Economics*, 88(9-10): 2099–2123.

Senik C. (2009), Direct evidence on income comparisons and their welfare effects, *Journal of Economic Behavior and Organization,* 72(1): 408-424.

Subramanian S.V. and I. Kawachi (2006), Whose health is affected by income inequality? A multilevel interaction analysis of contemporaneous and lagged effects of state income inequality on individual self-rated health in the United States*, Health and Place,* 12(2): 141-156.

Subramanian S.V., Kawachi I and B.P. Kennedy (2001), Does the state you live in make a difference? Multilevel analysis of self-rated health in the U.S*, Social Science and Medicine*, 53(1): 9–19.

Summerfield M., Freidin S., Hahn M., Ittak P., Li N., Macalalad N., Watson N., Wilkins R. and M. Wooden (2012), *'HILDA User Manual- Release 11",* Melbourne Institute of Applied Economic and Social Research, University of Melbourne.

Kahn R.S., Wise P.H., Kennedy B.P. and I. Kawachi (2000), State income inequality, household income, and maternal mental and physical health: cross sectional national survey, *British Medical Journal*, 321(7272): 1311–1315.

Kaplan G.A., Pamuk E.R., Lynch J.W., Cohen R.D. and J.L. Balfour (1996), Inequality in income and mortality in the United States: analysis of mortality and potential pathways, *British Medical Journal*, 312(7037): 999-1003.

Karlsson M., Nilsson T., Lyttkens C.H. and G. Leeson (2010), Income Inequality and health: importance of a cross-country perspective*, Social Science and Medicine*, 70(6): 875-885.

Kawachi I., Kennedy B.P., Lochner K. and D. Prothrow-Smith (1997), Social capital, income inequality and mortality*, American Journal of Public Health,* 87(9): 1491-1498.

Kennedy B.P., Kawachi I., Glass R. and D. Prothrow-Smith (1998), Income distribution, socioeconomic status and self-rated health: a US multilevel analysis, *British Medical Journal*, 317(7163): 917–921.

Krugman P. (1996), The spiral of inequality, *Mother Jones*, (November/ December): 44–49.

Lochner K., Pamuk E.R., Makuc D., Kennedy B.P and I. Kawachi (2001), State-level income inequality and individual mortality risk: a prospective multilevel study*, American Journal of Public Health*, 91(3): 385–391.

Lundberg O., Fritzell J., Yngwe M. Å. and M.L. Kölegård (2010), The potential power of social policy programmes: income redistribution, economic resources and health, *International Journal of Social Welfare,* 19(1): S2-S13.

Lynch J.W., Kaplan G.A., Pamuk E.R., Cohen R.D., Heck K.E., Balfour J.L. and I.H. Yen (1998), Income inequality and mortality in metropolitan areas of the United States. *American Journal of Public Health,* 88(7): 1074- 1080.

Mellor J.M. and J. Milyo (2002), Income inequality and individual health: evidence from the Current Population Survey, *Journal of Human Resources*, 37(3): 510–539.

Miller D. and C. Paxson (2006), Relative income, race, and mortality, *Journal of Health Economics,* 25(5): 979- 1003.

Rodgers G.B. (1979) Income and inequality as determinants of mortality: an international cross-section analysis, *Population Studies,* 33(2): 343-51.

Verbeek M. (2012), *A Guide to Modern Econometrics,* 4th edition, John Wiley & Sons.

Wagstaff A. and E. van Doorslaer (2000), Income inequality and health: what does the literature tell us?, *Annual Review of Public Health*, 21(1): 543–567.

Ware J., Snow K. and M. Kosinski (2000), SF-36 Health Survey: Manual and Interpretation Guide, *Quality Metric Incorporated*, Lincoln: RI.

Wilkinson R.G. (1996), *Unhealthy societies: The afflictions of inequality*, London: Routledge.

Wilkinson R.G. and K.E Pickett (2010), *Jämlikhetsanden: därför är mer jämlika samhällen nästan alltid bättre samhällen,* Stockholm: Karneval.

World Bank (2014), World Bank Development Indicators, http://data.worldbank.org/indicator, (2014-05-13).

Zhao Z. (2006), Income inequality, unequal health care access, and mortality in China, *Population and Development Review*, 32(3): 461–483.

Appendix

Table 3 Descriptive Statistics

Variable	Frequency	Min	Max	Mean	Std. Dev.
Gini	79 518	0.2049	0.4075	0.3110	0.0230
Income	79 519	0	525 146	37176.41	226639.59
Mean income_ref	78 980	18846.15	114827.6	37033.48	11505.24
Health	71 475	0.301	1	0.7612	0.1206