Empirical studies in the measurement of socio-economic inequality in health

S. Baeten

Baeten, S.

Empirical studies in the measurement of socio-economic inequality in health
Dissertation Erasmus University Rotterdam, the Netherlands

Cover: Photo by ant.photos. Design by Dennis Petri

Printing & Layout: Optima Grafische Communicatie, Rotterdam, the Netherlands

ISBN: 978-94-6169-384-6

Empirical studies in the Measurement of Socio-economic Inequality in Health

Empirische studies naar de meting van sociaaleconomische ongelijkheid in gezondheid

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de
rector magnificus

Prof.dr. H.G. Schmidt

en volgens besluit van het College voor Promoties.
De openbare verdediging zal plaatsvinden op

vrijdag 24 mei 2013 om 9.30 uur

door

Stefan Arthur Baeten
geboren te Nieuwegein

PROMOTIECOMMISSIE

Promotor:	Prof.dr. E.K.A. van Doorslaer
Overige leden:	Prof.dr. W.B.F. Brouwer
	Prof.dr. O.A. O'Donnell
	Prof.dr. M. Grimm
Copromotor:	Dr. T.G.M. Van Ourti

CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1

Introduction

Worldwide the demographic transition over the past 150 years has led to large increases in life expectancy. Cutler et al. (2006) link the early stages of this development to improved nutrition and economic growth. This stage is followed by large changes in public health, urbanization, sanitation and waste disposal. The final stage of this transition are attributed to better medical care, such as vaccinations, antibiotics and late 20th and early 21st century specialized medical treatments. Unfortunately, these improvements have not been equally impressive for all countries, nor for all within a single country. Social status is one of the major sources of this within country health variation. These inequalities will most likely remain in the center of attention for much longer, as trends show that for some indicators the gap is narrowing, but for other it is still increasing (Bleich et al., 2012). Explanations for these differences have been sought in medical care, resources, health behavior and social structures (2006). Nonetheless, much remains unknown and health inequalities by socio-economic status (SES) remain an intensively studied topic.

This interest in the SES-health gradient is by no means a purely academic one. The reduction of SES-health inequalities is an important goal of both governments and societies (Murray and Evans, 2003b). Especially Western governments explicitly aim to reduce disparities by income, education, work status and race (Mackenbach, 2002; Department of Health, 2003; Graham and Kelly, 2004; U.S. Department of Health and Human Services, 2010). International organizations are also much concerned with the issue. The World Health Organization views inequalities within and between countries as a result of political, social and economic choices that reflect the societal justice and development of a society (Commission on Social Determinants of Health (CSDH), 2008). The OECD stresses the unfairness and avoidability of such inequalities as the major reason for targeted reduction of SES-health inequalities by its member-states. The OECD therefore tries to support states in the development of good indicators of health inequalities (Looper and Lafortune, 2009). Finally the World Bank considers it its task to provide countries with reliable information on the severity of inequalities, so that governments can help the poor also take full advantage of health improvements (Gwatkin et al., 2007).

A discussion of the normative implications of the SES-health gradient, as claimed by these international organizations, is beyond the scope of both this introduction and thesis. In general, SES related health inequalities are regarded as inequitable (Braveman and Gruskin, 2003) and because of associated costs sometimes also as inefficient (LaVeist et al., 2009). This thesis does, however, contribute to the demand of these organizations for improved measurement and additional empirical evidence of the gradient. It presents a collection of studies related to the measurement of socio-economic health inequalities in a variety of settings. Each of the studies was motivated by one or more of the following three goals. A first goal was to contribute to empirical knowledge on what drives these inequalities by providing evidence for different countries, time periods and populations. A second goal was to illustrate how inequality measurement can be affected by certain

methodological problems and how these problems can be overcome. A third goal is to provide further insight in the consequences of inequalities.

In Chapter 2 I both deal with a methodological issue and provide new empirical evidence. Inequalities vary both over time and age. In that regard it is important to measure inequalities over the life cycle. So far these inequalities over the life cycle between high and low income groups are not well understood. Most current evidence suggests that at young ages the distribution is rather equal, since the majority of the population is in good health. At higher ages health deteriorates for both groups, but faster for low income groups than for rich income groups. However, around retirement age health seems to start to deteriorate faster for high income groups, diminishing the inequalities. This results in health levels that first diverge between rich and poor, before they converge again. Several mechanisms that could be responsible for this pattern have been put forward. Some of these suggest that biases in the measurement of health over the life cycle are responsible for this finding. In Chapter 2 I investigate one of these biases, selective dropout, in detail (Beckett, 2000; Lynch, 2003).

Health changes over time may seem to be measurable adequately if panel data is available, but even then issues arise related to dropout. When such attrition is related to health, the sample in one period is no longer comparable to that in the previous period. This is especially worrisome if different groups are compared, like SES groups, and the attrition is also different for those groups. Probably, the attrition resulting from premature death is the most important potential cause of bias in that respect. In Chapter 2 I also investigate the role of institutionalization, since individuals that move into nursing homes and other long-term facilities are usually no longer included in health surveys.

The problem is best illustrated using a simplified example. Think of two SES groups, each with different health levels and mortality rates. At a certain age, 20% of the low SES group is sick and 15% of the high SES group is. In both groups, all healthy individuals survive and none get ill, but the probability of surviving for the sick is different. Half of the sick from the low SES group die, whereas only 20% of the sick in the high SES group die. In the next period 11% of those who are still alive in the low SES group are sick and 12% of high SES survivors. So, even though the low SES group is initially disadvantaged and experiences worse health outcomes, the comparison of two cross-sections suggests their health improved compared to the high SES group. This phenomenon is called selective dropout and could explain the diverging and converging health inequalities between rich and poor groups over the life cycle.

In Chapter 2 I deal with this problem and illustrate the consequences for life cycle patterns of poor health for income groups in the Netherlands. I first confirm that life cycle patterns of bad health in the Netherlands are different for both income groups and follow the diverging and converging patterns. Next, I measure the degree of selective dropout in the Netherlands using a sample of more than 2.5 million Dutch. Finally, I propose a

correction using these results that reveals the life cycle patterns of low and high income groups in the Netherlands without selective dropout. I do find evidence for selective dropout by income and health. However, the effects on the life cycle patterns of health are small and do not explain the major share of the diverging and converging health patterns.

Chapter 3 deals less with methodological issues, but more to the gathering of good empirical evidence of socio-economic health inequalities. Like Chapter 2, Chapter 3 also deals with a time dimension in socio-economic health inequalities. It takes into account that socio-economic circumstances may have a delayed or continued effect on health. I investigate how levels of health in adulthood relate to SES during childhood and what pathways between child- and adulthood could explain this relationship. I investigate this using information on body mass index (BMI) using eight waves of the China Health and Nutrition Survey (CHNS). During the study period (1989-2009) China has undergone great changes in many ways. Of course there are economic developments that have been unprecedented, but also the way of life in China has changed a lot. Some of these latter changes, mainly those in diet and physical activity, have led to large increases in BMI and associated diseases. I relate these changes in life styles to childhood SES. Besides being able to identify long term effects of SES, this approach also reduces endogeneity in the SES-health gradient. Although the existence of the gradient is well established, the underlying causal mechanisms are less well understood. Groups with higher education or income may be more inclined or capable to invest in their own health. The effect could also be opposite. It could be that bad health would limit an individual's potential to achieve higher socio-economic status. Finally, third factors could have effects on both SES and health. For example genes could explain the socio-economic status someone achieves later in life as well as his or health. Some studies have used natural experiments as exogenous shocks in income (Ettner, 1996; Meer et al., 2003) or education (Oreopoulos, 2006; Kippersluis et al., 2009a) and analyzed those using instrumental variables regressions. It seems likely that all three explanations are relevant.

In Chapter 3 I use maternal education as an indicator of childhood SES and link this to BMI later in life. I assess what happens to the regression coefficients in different models in which possible explanatory pathways from childhood SES to BMI are added. The main focus is on the role of early adulthood health, adulthood SES and lifestyles. The results suggest that there is a childhood SES-adulthood BMI gradient in China, but only if it is differentiated by age. These differential effects imply that the negative marginal effect of a year of maternal education becomes larger if children get older and ultimately becomes positive. The SES-BMI gradient thus rotates with age. Both the gradient and its rotation disappear if adulthood SES is accounted for. This turns out to be the most important pathway from childhood SES to adulthood BMI. Interestingly, no gradient is found at the extreme ends of the BMI distribution. This eases concerns about the impact of SES on BMI, because the more extreme BMI levels are most detrimental to health.

In Chapter 4 I further investigate socio-economic health inequalities in China. Here I hypothesize a relationship between socio-economic health inequalities and exogenous economic developments. Few countries have experienced such rapid and large changes in their income distribution as China during the study period (1991-2006). The economic developments in China have had effects both on average income and the income distribution, but have also led to lower access to health services for the poor (Akin et al., 2005). If a link exists between income and health, such large changes in the income distribution can have an impact on the joint distribution of health and income (Contoyannis and Forster, 1999). What is more, if those who cannot improve their health as a result of increased income also have less access to good health services, this may amplify their health deprivation. Chapter 4 makes a methodological contribution by presenting a decomposition of the concentration index that shows that the large changes in the income distribution in China have been accompanied with increases in income related health inequality.

Empirically, a non-linear relationship between income, health and other exogenous health determinants is estimated and used within the decomposition to link changes in health and income. Although health is used as the dependent variable here, I claim no - nor do I have to for the purpose of this chapter - causality in this relationship. I find that the large changes in the income distribution in China have been accompanied with increases in income related health inequalities. The evidence from the decomposition shows that the most important driver of these increases is differential income growth for young and old. Additionally, increases in average income and income inequality contribute to the increased income related health inequalities. A further analysis shows that mainly older women move down in the income distribution, which is most likely a result of incomes not being protected by an adequate pension system.

Chapters 2 to 4 deal with the distribution of health *within* a country. Even larger variation in health may exist *between* countries, also related to economic well-being. The differences in average life expectancy between high- and low-income countries are already indicative of the socio-economic gradient in health inequalities between countries. These inequalities are not new and are most often described using a Preston curve (1975). However, these commonly use a single health outcome such as mortality or a disease specific outcome. For a complete picture of between country SES-health inequalities both dimensions of health, i.e. mortality and disability, are important. In that sense, life expectancy is an adequate measure of mortality, but pre-defined diseases will be insufficient to describe all variation in health related well-being. A general health-related quality of life (HrQoL) measure is more suitable to measure disability. Typically, HrQoL estimates and mortality data are combined in measures such as healthy life expectancy. In Chapter 5 I calculate such healthy life expectancies, compare inequalities between different income countries and propose two improvements to the HrQoL measurement in different countries.

Hitherto there is only a limited literature on this topic and by far the largest contribution has been made by the World Health Organization (2000, 2012b; Murray and Evans, 2003a). Its efforts have been twofold. First, it has focused on impact assessment of a wide variety of diseases in the global burden of disease (GBD) studies (Murray and Lopez, 1997). Second, it has used health surveys to obtain HrQoL estimates that are comparable across different countries. In Chapter 5 I contribute to the way in which these household surveys are analyzed. The first improvement I propose increases the variability in HrQoL estimates between individuals. The second one deals with an underestimation of HrQoL in especially richer countries. The effects of the first improvement on between country inequalities are limited, but can be very helpful in inequality measurement within countries. The second improvement is particularly useful in between country inequality measurement, because the impact of the improvement is larger in rich than in poor countries. Even without these improvements healthy life expectancy is much higher in the richest countries than in the poorest countries, 23.4 healthy life years to be precise. After the implementation of the improvements this inequality shows to be even larger, namely 26.3 healthy life years.

In chapters 2 to 5 I present estimates of health inequalities in a number of different settings and each chapter deals with one or more methodological challenges to estimating these inequalities. In Chapter 6 I present a way in which health inequality measurement can be used in evidence based policy making. The results of health inequality studies are most useful if they help reduce health inequalities. An important condition for this is that inequalities are measured on a regular basis. Such monitoring can show whether policies aimed at reducing inequalities are effective, but even then evidence based policy making is not straightforward (Macintyre, 2003). However, monitoring alone may be insufficient. First of all it does not provide a priori information available on the equality effects of proposed policies. And secondly, it does not help to weigh inequality and other policy effects. Therefore, in Chapter 6 I present three methods to evaluate health interventions simultaneously on their cost-effectiveness and their distributional impact. In health technology assessment interventions are traditionally only evaluated based on their health improvements and costs. This suggests a pure efficiency objective, because it aims at maximizing health output of the health systems with a given budget. Other criteria for reimbursement decisions, such as equity, necessity and the rarity of the disease, are generally evaluated separately and more qualitatively. The methods I propose aim at quantifying the equity implications of health interventions and make trade-offs with efficiency outcomes more explicit. I introduce these methods using an example from breast cancer control.

The first method I introduce calculates what health disparities exist between specific groups in the population before and after the introduction of a new health intervention. The second method values the remaining years of life of patients not only based

on HrQoL in those years, but also based on the already experienced level of health in the past. This allows for differential weights to be given to patients based on their age, diseases severity and co-morbidities. The third method includes additional criteria to efficiency and equity, such as the aforementioned necessity and rarity. In this method interventions are described based on all these criteria simultaneously. Subsequently, each intervention is evaluated using weights elicited from a discrete choice experiment. The three methods lead to different conclusions about the trade-off between equity and efficiency. They also vary in the possibilities that are left for policy makers to weigh different outcomes. In the first method these possibilities are more or less unaffected, whereas the second and third considerably limit these possibilities. However, each method provides tools to incorporate societal aversion to health inequalities in reimbursement decisions.

Chapter 7 discusses and concludes this thesis for each of its three goals of this paper. First I summarize and compare the empirical evidence in this thesis on the extent of socio-economic health inequalities. Then I discuss some of the major issues in SES health inequality measurement and how each chapter contributes to the literature on those issues. Finally, I summarize the policy implications of the evidence in this thesis on the first and second goal.

Chapter 2

The socioeconomic health gradient across the life cycle: what role for selective mortality and institutionalization?

In collaboration with: Tom Van Ourti and Eddy van Doorslaer

ABSTRACT

Background

Many studies have documented the now fairly stylized fact that health profiles by income differ across the age distribution: the health gap between rich and poor tends to widen until about age 50 and then declines at higher ages. It has been suggested that selective mortality and institutionalization could be important factors driving this convergence at higher ages.

Data

We use eight waves of a health survey linked to four registries (on mortality, hospitalizations, (municipal) residence status and taxable incomes) to test this hypothesis.

Method

We construct life cycle profiles of health by birth year/gender/income groups and obtain precise estimates of individual probabilities of mortality and institutionalization using a seven year observation period for more than 2.5 million individuals. We then generate selection corrected health profiles using an inverse probability weighting procedure.

Results and conclusion

We find that attrition is indeed not random: older, poorer and unhealthier individuals are significantly less likely to survive the next year and not to be admitted to an institution. But while these selection effects are very significant, they are not very large. Therefore, we reject the hypothesis that selective dropout is an important determinant of the differential health trajectories by income over the life course in the Netherlands.

2.1. INTRODUCTION

Recently, the evolution of health disparities by socioeconomic status (SES) over the life cycle has been studied by scholars from a variety of disciplines (Beckett, 2000; Lynch, 2003; Case and Deaton, 2005; House et al., 2005; Sacker et al., 2005; Smith, 2005; Herd, 2006; Mirowsky and Ross, 2008). In this literature, two hypotheses are generally put forward on how this SES health gradient evolves over the life cycle: *cumulative advantage* and *age-as-leveler*. Both hypotheses assume that health deteriorates with age up to late middle age and that the rate of deterioration is steeper for low than high SES individuals. However, the hypotheses differ in how to assume health inequalities develop thereupon. Proponents of cumulative advantage argue that health differences between SES groups widen until the late stages of life, while age-as-leveler states that health disparities start converging from late middle age onwards. A potential explanation for such a converging trend is that the biological determinants of ageing (and thus health deterioration) start dominating the influence of SES after late-middle age (Herd, 2006), with most prominence in the literature so far being given to cohort effects (Lynch, 2003) and selective mortality (Beckett, 2000).

In this paper we focus on the role of selective mortality and institutionalization – mortality differences and differences in institutionalization rates between high and low SES groups – in shaping the life course profile of the SES health gradient. Selective mortality seems a likely candidate as an explanation for the converging SES health gradient after late-middle ages as relative mortality differences have been found to follow a similar life course pattern. For instance, relative mortality differences between high and low SES individuals have been found to peak around ages 50-60 and to fall again at higher ages (Kippersluis et al., 2010). When a high SES individual is more likely to survive than a low SES individual – even when both are equally (un)healthy – then at older ages, the high SES group will include on average a greater share of unhealthy individuals than the low SES group. Similarly, if the low SES and unhealthy are more likely to move into an institution than their equally unhealthy high SES counterparts, the SES health gradient would start falling around those ages when individuals move to institutions. Both types of selective dropout may contribute to the apparent narrowing of the health gradient above a certain age in cross-sectional evidence.

In this paper we aim to add to the existing literature on this phenomenon in five respects. We are the first to study the importance of selective mortality for the life cycle profile of the SES health gradient in the Netherlands. With the exception of one Canadian (Prus, 2007) and two UK studies (Sacker et al., 2005; McMunn et al., 2009), all other studies have been based on US data. Second, our data allows for an analysis of selection effects due to residential long term care. Our survey, like most others, also only includes private households and excludes individuals living in institutions such as nursing homes. Other studies, such as Beckett (2000), have typically studied the effect of *overall* dropout (includ-

ing mortality and residential long-term care use), but our data allows for the identification of individuals that have died or have moved into a residential institution. The Netherlands is a particularly interesting country to study this phenomenon since Dutch elderly have relatively higher utilization rates of residential long term care than elderly in other OECD countries (Colombo et al., 2011): in 2009, almost 10 percent of the 65+ population was living in a residential long term care institution (Statistics Netherlands, 2011).

Third, our analysis has substantial statistical power to study selective attrition as we use a set of repeated cross-sectional surveys linked to administrative data covering the period 1998-2005. The use of these administrative data allows for a rather precise identification of the mortality and institutionalization probabilities due to the much larger sample size (i.e. over 2.5 million observations) than what is typically available in surveys. The administrative data also provide us (after linking with the surveys) with an adequate period of follow-up for the survey respondents.

Fourth, we propose a new method to control for selective attrition by combining an approach examining health levels at an aggregated group level (Deaton and Paxson, 1998b) with inverse probability reweighting at the individual level (Jones et al., 2006). This combination of methods has not featured in the literature on selective attrition and the SES health gradient before, but has been used in related fields (e.g. Kippersluis et al., 2009b). The aggregated approach consists of transforming the set of repeated cross-sections into a panel dataset of age-groups, thereby allowing for the identification of life cycle effects. Inverse probability weights (IPWs) are used to correct for selective attrition. They are derived from an individual's predicted mortality/institutionalization conditional on past individual characteristics such as age, SES and health status. It thus, and unlike the imputation method of Beckett (2000), considers the experiences of *both* survivors/non-attritors and decedents/attritors.

Finally, we allow attrition to depend on the interaction between prior SES and prior health status. Again, this is a crucial feature and a major advantage of our approach, because selective mortality will matter most for the SES health gradient if and when the association between mortality and ill-health differs *across* SES groups. Instead, all prior studies have relied simply on interactions between age and SES in explaining differential health trajectories.

2.2. PREVIOUS LITERATURE

Beckett (2000) as the first to analyze the combined influence of selective mortality, institutionalization and other sources of sample attrition in the US. By exploiting changes between the 1982-1984 and 1992 waves of the NHANES I Epidemiologic Follow-up Study she was able to isolate the overall effect of attrition on the SES health gradient. Her findings confirmed that health by SES differences widen till late middle age and start

converging thereafter, but she also found that the convergence was not due to overall sample attrition. Noymer (2001) criticized her approach because it relies on the assumption that the counterfactual health levels of the deceased may be overestimated because they were imputed from the experience of the survivors only. Another criticism was that the counterfactual health levels differed by age, SES and the interaction between the two, but that the interaction between initial health and initial SES was not accounted for. The latter interaction seems essential in the context of selective mortality as it would allow for different mortality effects of ill-health across SES groups. It is therefore not very surprising that Beckett (2000) found no effect of selective mortality on converging SES health differences at older ages. In later work, Beckett and Elliott (2001) have partially addressed both concerns by estimating mortality models on the sample of survivors and deceased and by conditioning on initial health status and its interaction with initial SES, but they did not use these models to analyze the effect of selective mortality on the SES health gradient.

Another influential approach – mainly credited to Lynch (2003) – consists of analyzing SES-differences in the life cycle profile of health using a random coefficients model. It has generally found evidence in favor of converging SES health differences at older ages being driven to some extent by selective mortality (Sacker et al., 2005; Herd, 2006; Mirowsky and Ross, 2008). A crucial feature of this approach concerns the splitting of overall health variation into variation between and within individuals. The life cycle behavior of the SES health gradient is then derived from the within individual variation only and, in contrast to Beckett's (2000) approach, one does not have to impute counterfactual health levels for the deceased (or other individuals that dropped out of the panel survey). The main intuition is that purging within-individual health variation from between-variation should eliminate the effect of selective mortality (and attrition), but this will only hold as long as the deceased/attritors are "missing at random". Moreover, the random coefficients approach only implicitly reveals the effect of selective mortality (and attrition) on the life cycle profile of the SES health gradient, and is thus not a natural choice for our purposes.

We discuss our aggregated group IPW-approach in detail in sections 2.3 and 2.4, but for transparency we highlight the crucial differences with Beckett (2000) and Lynch (2003) here as well. It differs from the random coefficient models to the extent that we model selective mortality and dropout due to institutionalization explicitly. It also differs from the imputation approach of Beckett since our correction is based on the experiences of both survivors and decedents, and we also condition on the interaction between initial SES and initial health status. A similarity with the earlier mentioned approaches is that we will impose that observations are "missing at random" conditional on the variables used to explain the IPWs. This assumption has recently been criticized by Petrie et al. (2011). They argue that "missing at random" contradicts the idea that death itself is an indication that health deteriorates more rapidly for the deceased than for the survivors.

Hence, using IPWs will lead to a lower bound on health disparities between rich and poor, and this should be kept in mind when interpreting the results of our aggregated group-IPW approach. Petrie et al. (2011) propose to impute the health levels of deceased individuals with a(n absolute) zero health level which is the exact opposite of "missing at random" and will inevitably give rise to an upper bound on SES health differences. While we agree with the Petrie et al. (2011) criticism, we believe it is less relevant here since the construction of our IPWs allows for dynamics in the effect of the explanatory variables.

While the importance of selective mortality and institutionalization for the SES health gradient has not been analyzed for the Netherlands before, van Kippersluis et al. (2010) do provide some evidence for the age-as-leveler hypothesis using self-reported health and different indicators for SES, including current income, education and occupation. They show that the health deterioration for low SES individuals slows down and even reverses after the age of 55. While the underlying mechanisms leading to this improved health situation have not been well understood, a similar pattern has been reported for the US by Smith (2004). Van Kippersluis et al. (2010) also find that cohort effects in the SES health gradient in the Netherlands are absent and this has also been confirmed with other Dutch survey data (Kippersluis et al., 2009b).

2.3. THE LIFE CYCLE PROFILE OF HEALTH ACROSS INCOME GROUPS WITHOUT CORRECTION FOR SELECTIVE DROPOUT

Our main goal is to understand the importance of selective mortality and institutionalization for converging health disparities between high and low SES groups after age 55 in the Netherlands. We start off by constructing baseline life cycle ill-health profiles for different SES groups without correcting for selective mortality and institutionalization. We use an aggregated group-approach as used in e.g. van Kippersluis et al. (2010) to disentangle cohort and life cycle effects from a set of repeated cross-section surveys. It requires dividing each cross-section into subgroups based on gender, birth year and level of SES. By following these gender/birth-year/SES groups over time, cohort effects can be disentangled from life cycle effects for each gender/birth-year/SES group.

We use data from representative samples of non-institutionalized Dutch individuals taken from eight annual cross-sectional Surveys of Living Conditions held between 1998 and 2005 (SLC hereafter).[1] We restrict our analysis to individuals aged between 37 and 84 in 1998 and follow them over time. Since we use a set of cross-sections, we observe different individuals in each cross-section, but we make sure to adjust the specified age ranges, i.e. 37-84 in 1998, …, until 44-91 in 2005. The restriction at age 84 is the result of a

1. We restrict attention to 1998-2005 since at the time of the analysis the linkages to the Dutch administrative data – needed to illustrate the effect of selective dropout – were only available for 1998-2005.

trade-off between small cell sizes at advanced ages and the need to include sufficiently old individuals to study the effect of institutionalization (and mortality). The restriction at the lower end is because very few individuals die or become institutionalized at younger ages. Earlier work (see also section 2.2) confirms that most of the diverging and converging of SES-specific life cycle ill-health profiles occurs within this age range. The actual size of the sample used to construct the ill-health profiles amounts to almost 128,000 observations, and an important advantage for our purposes is SLC's linkage to administrative data on (a) cause of death, (b) taxable income, (c) hospital admissions and (d) residential status, which enables us to account for selective mortality and institutionalization. We discuss these links to administrative data in section 2.4(b) and 2.4(c).

Our procedure to construct baseline life cycle ill-health profiles for different SES-groups consists of 4 steps. First, we categorize each cross section of the SLC in groups defined by SES and year of birth. Since ill-health experiences differ greatly between sexes, we do this separately for males and females. We use current disposable after-tax household income as a measure of SES, and correct for household composition using the equivalence scale of Statistics Netherlands (Siermann et al., 2004) and express in 1997 prices. [2] We distinguish into four income quartiles separately for each cross-section to minimize the risk of small cell sizes.[3] We further subdivide each of the income quartiles into groups based on birth-year. We use sufficiently wide intervals to avoid small cell sizes, i.e. we define 12 birth-year intervals: born 1914-1917, 1918-1921, to 1958-1961. Summary statistics of the individual data underlying the first and fourth income quartile are provided in Table 2.1 in the appendix.[4]

Second, we calculate the average level of ill-health of each gender/birth-year/income group in each cross-section as the proportion of respondents reporting to be in bad

2. We prefer to use income over other indicators of SES such as occupation and education since (a) such alternative indicators are not reported in our administrative data sources and (b) earlier evidence shows that Dutch ill-health life cycle profiles across SES groups are consistent across occupation, education and current income (Kippersluis et al., 2010).

3. An additional advantage of cross-section specific income quartiles – contrary to pooled-sample-quartiles – is that individuals from earlier cross-sections are not more/less likely to end up in lower quartiles due to life cycle profiles of incomes and average income growth between 1998 and 2005.

4. Table 2.1 does not provide summary statistics for the middle two income quartiles since we only report results based on the first and fourth income quartile throughout the paper (see also below).

TABLE 2.1: Summary statistics of the individual data underlying quartile 1 and 4 in the SLC sample

	Income quartile		1998	1999	2000	2001	2002	2003	2004	2005
Males	1	% bad health	44%	46%	45%	42%	44%	44%	44%	49%
		mean income	€ 10,845	€ 9,926	€ 10,372	€ 13,487	€ 12,935	€ 13,097	€ 12,821	€ 13,079
		median income	€ 11,558	€ 10,761	€ 11,268	€ 14,145	€ 13,619	€ 13,806	€ 13,506	€ 13,682
		iqr income	3,678	3,831	3,775	3,915	4,098	4,043	3,856	3,886
		N	4,203	1,912	1,656	1,038	1,186	1,006	1,003	434
	4	% bad health	13%	14%	13%	16%	16%	17%	18%	18%
		mean income	€ 45,729	€ 43,977	€ 47,481	€ 50,540	€ 48,718	€ 49,355	€ 47,337	€ 51,866
		median income	€ 38,302	€ 37,242	€ 38,691	€ 44,594	€ 43,685	€ 44,213	€ 43,426	€ 43,769
		iqr income	13,442	12,223	12,969	13,173	13,142	13,189	11,715	15,574
		N	6,033	2,784	2,338	1,320	1,464	1,334	1,362	598
Females	1	% bad health	45%	49%	46%	45%	46%	47%	48%	52%
		mean income	€ 10,262	€ 9,794	€ 10,239	€ 12,972	€ 12,459	€ 12,762	€ 12,610	€ 12,460
		median income	€ 10,668	€ 10,141	€ 10,445	€ 13,228	€ 12,691	€ 13,078	€ 13,035	€ 12,610
		iqr income	3,818	3,707	3,753	4,400	4,291	4,095	3,999	3,909
		N	6,832	3,096	2,600	1,503	1,689	1,587	1,544	665
	4	% bad health	17%	20%	19%	21%	19%	20%	22%	19%
		mean income	€ 45,545	€ 43,832	€ 45,824	€ 51,487	€ 50,522	€ 49,564	€ 48,380	€ 51,094
		median income	€ 38,260	€ 37,247	€ 39,036	€ 45,483	€ 44,524	€ 44,485	€ 43,793	€ 44,846
		iqr income	13,373	12,422	13,385	14,756	14,411	13,155	13,253	15,709
		N	5,005	2,224	1,919	1,220	1,409	1,259	1,185	498

Note: iqr: inter quartile range

health as derived from the self-reported health question "How good is your health in general?" The bottom three[5] response categories are classified as bad health.[6]

Third, the previous steps provide us with a panel dataset. For each gender/income category, we have 96 observations of average ill-health (i.e. 8 cross-sections times 12 birth-year intervals). Variation over time *within* birth year groups can be used to estimate life cycle ill-health profiles, while variation *between* birth year groups allows identification of

5. The phrasing of the two worst response categories changed from 2000 to 2001. In the earlier waves these were "sometimes good, sometimes bad" and "bad" and later they were "bad" and "very bad" However, the third worst response category was described in all waves as "it's ok".

6. The findings reported in this paper are robust to alternative definitions of bad health including (a) measuring bad health as the worst response category only and (b) using the interval regression approach of van Doorslaer and Jones (2003) SAH might also be prone to response heterogeneity by age (Bago d'Uva et al., 2008a, 2008b). When older individuals have lower expectations of their function abilities, and when lower income groups more rapidly adjust their expectations than higher income groups, reporting heterogeneity could be related to converging health disparities between rich and poor at older ages. Unfortunately, lack of vignettes data do not allow to analyze this in more detail.

cohort effects by year of birth. This will only hold if period effects are unimportant for the measurement of ill-health. In step 3, we run a separate OLS regression on the 96 birth year/gender/income groups and model average ill-health as a function of birth year and age. By using dummies to represent birth year and age, one can separately identify life cycle and birth year effects in a very flexible way. We use 12 dummies for birth year and 26 dummies for age.[7] It turned out that the birth year dummies were jointly insignificant in seven of the 8 OLS regressions (i.e. one per income quartile, separately for males and females).[8] This is in line with earlier work by van Kippersluis et al. (2010) who report no or very small cohort effects in the evolution of the income health gradient in the Netherlands using a much longer follow up period, i.e. between 1983 and 2000. Therefore, we base the remainder of the analysis on OLS regressions of average ill-health on the set of age dummies only.

Fourth, we show the OLS-predicted average ill-health levels across the life cycle in a graph for each gender/income classification. We use a second degree Gaussian weighting kernel to smooth these lines over the life cycle. To avoid an overload of information, we only present the ill-health life cycle profiles for the lowest and highest income quartile.[9]

Our baseline scenario against which the effect of selective mortality and institutionalization will be assessed is presented in Figure 2.1. It shows the ill-health life cycle profiles of the lowest and highest income quartile for both genders. We find *prima facie* evidence in favor of the age-as-leveler hypothesis: average health decreases for both income quartiles and genders from age 40 onwards, but more steeply for the poorest male quartile. Around ages 50-55, health of the lowest income quartile improves until around ages 60-65 while health levels of the highest income quartile evolve at a more constant rate, and the converging trend is more pronounced for males than females. Nonetheless, the gap also narrows for women, this narrowing seems to take place around age 60. Differences between the profiles for men and women were less striking

7. Each birth year dummy corresponds to an interval of 4 birth years. The age dummies correspond to the (mid-interval) age of each birth year interval in 2 consecutive cross-sections. We have age dummies for the ages 39, 41, 43, …, 89. For example, the youngest birth year interval (those born between 1958 and 1961) is between 37 and 40 in the SLC of 1998 and between 38 and 41 in the SLC of 1991, hence the mid-interval age is 39. Similarly, the mid-interval age of the youngest birth year interval equals 41 in the SLC of 2000 and 2001. Consecutive age dummies are similarly defined.

8. Since our age dummies are 2 years wide and our birth year dummies 4 years wide, we only observe gender-income groups of the same age that were born within a birth year range of 8 years. Any interpretation of differences in birth year effects that are beyond these 8 years, is based on an extrapolation of our estimates under a parallel trend assumption. Our data do not allow to test the likelihood of this assumption.

9. The ill-health life cycle profiles of the middle income quartiles in the baseline scenario – but also in the scenarios with selectivity correction – reveal patterns in between those of the lowest and highest quartile reported in Figure 2.1.

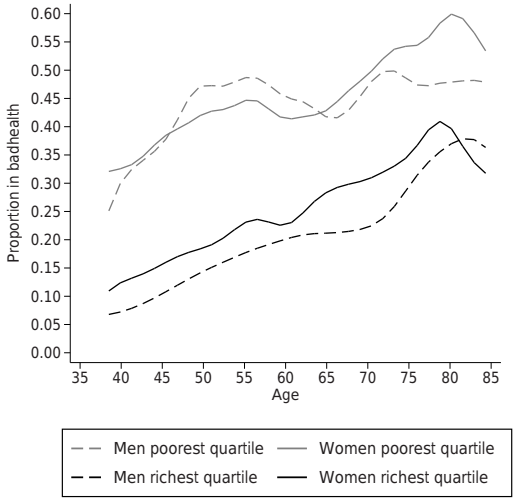


FIGURE 2.1: Baseline life cycle patterns of poor health for men and women of the richest and poorest quartiles in the Netherlands (1998-2005)

with a different definition of ill-health. What is more, we are not aware of such explicit differences in other countries either. Although, this is a striking result, it is beyond the scope of this study to explain it. At more advanced ages, the profiles are less smooth due to fewer observations in the birth year intervals. We also observe that average health among the highest income quartile is better for men than women, while the life cycle profiles of the poorest income quartiles cross twice. Our findings of first diverging and then converging differences between low and high income quartile's health levels, and the observation that these patterns are more pronounced for males, are in line with the literature reviewed in section 2.2, and earlier studies for the Netherlands (Kippersluis et al., 2010) and the US (Smith, 2004).

2.4. THE LIFE CYCLE PROFILE OF HEALTH ACROSS INCOME GROUPS AFTER CORRECTION FOR SELECTIVE MORTALITY AND INSTITUTIONALIZATION

In this section, we demonstrate and discuss the effect of correcting the ill-health life cycle profiles in Figure 2.1 for selective mortality and institutionalization. We do this by combining the grouping approach explained in section 2.3 with IPWs (Jones et al., 2006). The section consists of three parts. First, we explain how the IPWs can be combined with our grouping approach. Next we derive the IPWs for the case of selective mortality and for the case of institutionalization. We end by discussing the empirical effect of both corrections on the health life cycle profiles in Figure 2.1.

(a) Upgrading our aggregated group approach with IPWs

The aggregated group approach allows to understand how average ill-health of each gender/birth year/income group changes when its members grow 7 years older (i.e. from 1998 to 2005), but it does not disentangle whether the evolution of each group's average ill-health is driven by changes in its composition (e.g. selective mortality and institutionalization) and/or by changes in the health status of its members. The IPWs correct for these compositional changes during 7 years and thus enable us to remove the effect of selective dropout from the ill-health life cycle profiles.[10]

We explain the construction of the IPWs in the next section. Here, we discuss the correction of the ill-health life cycle profiles for selective dropout. At this stage, it is sufficient to know that the IPW of an observation in year *t* measures the inverse of this observation's probability to belong to the age range that was 37 to 84 in 1998. Hence, in the first cross-section of 1998, all observations have an IPW equal to 1 since those in the age range 37-84 constitute the group that we follow over time. From 1999 to 2005, observations' IPWs will deviate from 1 when there is selective dropout (in the unlikely case that all dropout is random, all IPWs remain equal to 1). For example, an observation in the age range 42-89 in year 2003 with an IPW larger than one has an above average probability to drop out between 1998 and 2003, and vice versa for IPWs smaller than one.

The correction for selective dropout in our aggregated group approach is then obtained by including these IPWs into the stepwise approach explained in section 2.3. In the first step, we subdivide – for each year of the SLC – those observations in the 37-84 age range in 1998 along the income dimension by constructing income quartiles that are equal in size in terms of the IPWs, i.e. the sum of the IPWs in each income quartile should be the same.[11] We further subdivide into groups based on birth-years. Next, we calculate the IPW-weighted average levels of ill-health of each gender/birth-year/income group to incorporate any selective dropout within each of these groups. The intuition behind this is that groups of individuals that are *initially* similar to those that drop out are inflated in order to keep the original size and composition of the gender/birth year/income group in 1998. The third and fourth step are similar as in section 2.3. The resulting ill-health life cycle profiles reflect the evolution of the (hypothetical) population without selective mortality and without selective dropout.

10. Selective dropout that might have occurred before we observe our sample in 1998 is not addressed by the IPW's, but more generally can only be addressed when pre-1998 data would be available.

11. In other words, this accounts for the fact that selective dropout might affect the relative size of income quartiles. For example, when low income individuals are more likely to drop out, the lowest income quartile after accounting for the IPWs will include fewer observations but will represent one fourth of the group that would occur without selective dropout.

(b) Calculating IPWs – selective mortality

The construction of IPWs is rather straightforward when applied to panel data, but we intend to apply the concept of IPWs to the cross-sectional SLCs. For readers unfamiliar with their use, we first explain their derivation with panel data (consult Wooldridge, 2002a section 17.7.3 for an overview) before explaining the procedure that we adopted for our set of repeated cross-sections.

When the same individual is observed repeatedly in a longitudinal setting, one can estimate selective mortality (or dropout) using a binary dependent variable model. The traditional IPW-approach models the probability that each individual present in the first wave (t=1) survives up to a future wave t as a function of explanatory variables in the first wave (t=1). A similar model (on all individuals in wave *1*) is run separately for each future wave t=2,…,T, and IPWs in wave t are obtained as the inverse of the predicted probabilities of survival resulting from the binary dependent variable model for wave t. As in the imputation approach of Beckett (2000) and the random coefficient models (Lynch, 2003), this IPW-approach thus assumes that the unexplained part of survival is 'missing at random', but – in contrast to the imputation approach – it uses variation from *both* decedents and survivors; and unlike the random coefficient model, it allows for a more flexible and less parametric correction for selective mortality. An alternative approach – discussed in Wooldridge (2002b) – is to model yearly survival probabilities when selective dropout is an absorbing state, which obviously holds for our case of selective survival. In practice one regresses survival in wave t as a function of explanatory variables in the previous wave t-*1*, conditional on those individuals that have survived up to the year t-*1*. In other words, this is equivalent to a procedure of 'cumulative' binary dependent variable models that first estimates survival from year *1* to *2*, next from year *2* to *3*, and so on until year T. The IPWs for year t are then obtained as the product of the IPWs of the binary dependent variable models for year *2* until year t. The main advantage of this approach is that the dynamics in the explanatory variables are used to predict survival. This should lead to a better explanation of actual survival, and thus makes the assumption of 'missing at random' more plausible.

Since the SLC consists of a set of repeated cross-sections, we cannot observe which individuals in the cross-section of year t=1999,…,2005 were unlikely to be present in the previous year, not to mention 1998. Without adding external information to these cross-sections, it is thus impossible to account for selective mortality using either of the two IPW approaches. Instead, we add external information obtained from three linked administrative data sources, i.e. (1) the cause of death registry (CoD), (2) the regional income survey (RIS) that includes taxable income records, (3) the municipal population register (MPR) and (4) the hospital admissions registration (HA).

The CoD registry, which provides date of death (if any) for all Dutch citizens, is used to construct a binary indicator of whether the individual was alive at 31 December of year

t. While this informs us whether an individual was alive for every year *t*=1999,…,2005, it does not identify individuals that are not included in the SLC due to mortality between 1998 and the year of observation in the SLC. We thus need more external information in order to model selective survival. The MPR registry is used in a similar way to identify individuals that enter a non-private household, which in most cases means a nursing home (more information is provided in section 2.4(c)).

The SLC is also linked to the administrative panel dataset RIS containing the annual income and tax records for a representative sample of one third of the Dutch population (more than 5 million individuals). As the RIS is a panel, we can apply the IPW approach and the large sample enables very precise estimation of the IPWs. We use all individuals in the RIS in 1998 in the relevant age range (2,521,122 observations) and model their survival until 2005. Note that restricting the analysis to individuals that feature in both the SLC and the RIS would not work since selective survival between 1998 and the year when the individual is observed in the SLC would be neglected. Instead, using all individuals in the RIS of 1998 is justified since this is a representative sample and the IPW approach imposes the 'missing-at-random' assumption. The latter assumption allows to assign IPWs calculated from the full RIS on all individuals included in the SLC based on the values of the explanatory variables included in the survival prediction model.

We argued in the introduction and section 2.2 that a correction for selective mortality should preferably account for differential survival patterns for those with similar health, but different income levels. Indeed, it seems likely that a process of selective mortality will be unimportant for the shape of the ill-health life cycle profiles when differences in survival rates across income groups do not depend on health status (or vice versa).[12] In the baseline computations in section2.3, ill-health was measured using a self-reported variable and ideally we would include the prior level of the same self-reported health variable in the prediction models underlying the IPWs.[13] However, the RIS and CoD do not report health information, and we only observe current (not past) self-reported health status in the SLC. Hence, we exploit the linkage between RIS and HA. While HA neither includes the same self-reported health variable, it records all hospital admissions in the Netherlands for all years *t*=1998,…,2005. We use prior hospitalization (defined as a dummy that equals 1 when there was at least one overnight stay in the hospital during

12. In our procedure explained in section 2.3, this would matter for the construction (and thus the composition) of the income quartiles; but as soon as one corrects for this income effect, the average health levels per income quartile will not change since the survival effect of health is similar across income groups.

13. We do not have to worry about potential endogeneity of prior health status, as we are not after the causal impact of prior health on survival, but just intend to correct the life cycle profiles of ill-health for selective mortality.

the previous calendar year) as a measure of prior health in the prediction models for the IPWs.[14]

We thus estimate a set of binary prediction models on the RIS exploiting the linkages with CoD and HA. Since selective survival is an absorbing state and since we believe that the dynamics in the interaction between lagged health and income (and more generally all explanatory variables) are essential to model selective mortality, we have modeled yearly survival probabilities using the approach of 'cumulative' binary choice models. We have estimated 14 logit models –for males and females in each t=1999,…,2005 – for survival up to year t as a function of age in 1998, and income and hospitalization in year t-1.[15] Age is included using one-year age dummies. Contrary to section 2.3, income is now defined in absolute rather than relative terms because an absolute income concept remains invariant over time.[16] If we were to use a relative income concept, survival from year t-1 to t would depend on the income quartiles in year t-1, but also on potential bias in these income quartiles arising from selective mortality between 1998 and t-1. This would require a correction of the income quartiles in year t-1. Instead an absolute income concept remains invariant over time and is therefore not prone to this additional form of selectivity bias. We define four income groups (corrected for household composition and expressed in 1997 prices), i.e. €0-€12,500; €12,500-€17,500; €17,500-€30,000 and €30,000 and higher. We initially allowed for all one-way interactions between these explanatory variables, but removed the interactions between age and hospitalization since these proved jointly insignificant. We do include the interactions between each income group and hospitalization, but for the interaction between age and income we combine some age dummies to allow for the limited variation in mortality at relatively young ages: we end up using younger than 64, 65-69, 70-74 and one-year age dummies above 75. We also tested for the two-way interaction between age, income and hospitalization, but found – in line with the insignificance of the interaction between age and hospitalization – no statistical evidence for this interaction. The IPWs for each year t are obtained as the product of the IPWs of the logit models for years 2 to t. It is important to emphasize that the precision of the resulting IPWs is high owing to the large sample size.

The final step projects the IPWs estimated for every individual in the RIS onto the individuals in the SLC. This would be straightforward if every individual in the SLC would

14. In the SLC, those having no overnight stay show a more favorable distribution of self-reported health than those having at least one overnight hospital stay.

15. Note that these models do not differentiate between age and birth year effects. This is unimportant since we are ultimately interested in the IPWs only (not so much in the underlying effect of regressors) and since we found no evidence for birth year effects in the 'uncorrected' ill-health life cycle profiles.

16. The income definition in the SLC and the RIS is identical, i.e. current disposable after-tax household income.

also be included in the RIS. However, as the RIS is a representative sample of around one third of the total Dutch population, while the SLC is a much smaller representative sample of the same Dutch population, we end up with (around) two thirds of SLC respondents without an exact link to the RIS. Therefore, an additional procedure was required to assign the IPWs from RIS to *all* individuals in the SLC.[17,18] For every gender and year t=1999,…,2005, we ran an auxiliary OLS regression (on all RIS individuals that survived until year t) of the predicted probability to survive up to year t (obtained from the logit models) upon age, hospitalization in year t-1, and income in year t, and impose the same set of one-way interactions. We use current income as we cannot retrieve previous year's income for the individuals that are in the SLC but not in the RIS. We also apply a logit transformation to the dependent variable before running the auxiliary OLS regressions since predicted probabilities always lie between 0 and 1. The estimated coefficients are then applied to all individuals in the SLC, and (after retransforming) provide us with the IPWs used to correct the ill-health life cycle profiles in Figure 2.1 for selective mortality.[19]

Table 2.2 provides summary statistics of the RIS data (with linkage to CoD and HA). We see that after one year (see column 1998) 1.18 percent of the 1,249,737 men and 0.72 percent of the 1,271,385 women of the 1998 RIS sample has dropped out due to mortality and this accumulates to 9.14 percent and 6.47 percent respectively in 2005.[20] We have applied the entire procedure described in this section on the RIS data. In order to avoid an overload of tables, we do not report the full estimates obtained for the 14 'cumulative' logit models and the 14 auxiliary OLS regressions, but we do discuss the most important findings and provide a graphical presentation of the resulting IPWs in Figure 2.2. All aux-

17. Even without an exact link, the 'missing at random' assumption would allow assigning the IPWs based on the values of the explanatory variables in the logit models only (i.e. neglecting the unexplained part). This does not work in our case since we only observe current income – and not the entire sequence of income levels since 1998 – for individuals in the SLC, but not the RIS.

18. We did a sensitivity test of the selective-mortality-corrected ill-health life cycle profiles using only those individuals that are included in the RIS and the SLC. For these individuals we can directly use the IPWs from RIS without the additional procedure (and without imposing its associated assumptions), but this comes at the price of being left with a sample that has (too) little power to analyze differences in ill-health life cycle profiles between income quartiles. The resulting estimates of the ill-health life cycle profiles are indeed less precise, but generally confirm the results based on the IPWs obtained from the additional procedure.

19. Note that this procedure will yield different IPWs from a procedure where the IPWs are directly obtained from modeling actual survival as a function of age, hospitalization in year t-1 and income in year t. In both approaches, we have a potential endogeneity bias of selective mortality driving current income (which we cannot address with our data), but this is less severe in our procedure since it also accounts for the (absence of) correlation between incomes in year t and t-1 (i.e. the income dynamics), and income in *year* t-1 and the other regressors.

20. These cumulative percentages are based on the cumulative number of deaths in the sample up until year t divided by the size of the RIS sample in 1998.

TABLE 2.2: Summary statistics of the RIS sample

		1998	1999	2000	2001	2002	2003	2004	2005
Males	Share of total sample* dropped out due to:								
	Death within next year	1.18%	1.31%	1.40%	1.53%	1.61%	1.67%	1.75%	
	Moving to an institution	0.08%	0.09%	0.09%	0.10%	0.11%	0.17%	0.19%	
	Average age	52.7	53.5	54.3	55.1	55.8	56.6	57.4	
	Share in income group:								
	€0-€12,500	24%	23%	20%	16%	14%	14%	15%	20%
	€12,500-€17,500	31%	31%	30%	28%	27%	28%	28%	32%
	€17,500-€30,000	31%	32%	33%	34%	35%	35%	34%	31%
	€30,000 and higher	14%	15%	16%	22%	24%	23%	23%	17%
	Share hospitalized	9%	9%	9%	9%	10%	11%	12%	
Females	Share of total sample* dropped out due to:								
	Death within next year	0.72%	0.83%	0.94%	1.07%	1.19%	1.26%	1.36%	
	Moving to an institution	0.18%	0.18%	0.20%	0.22%	0.25%	0.32%	0.37%	
	Average age	53.1	53.9	54.8	55.7	56.5	57.3	58.2	
	Share in income group:								
	€0-€12,500	28%	26%	24%	19%	17%	17%	17%	24%
	€12,500-€17,500	30%	30%	29%	28%	28%	28%	29%	31%
	€17,500-€30,000	29%	30%	31%	32%	33%	33%	33%	29%
	€30,000 and higher	12%	13%	15%	20%	22%	22%	21%	16%
	Share hospitalized	10%	10%	10%	10%	11%	12%	13%	

Note: *: The total sample at the baseline year 1998 consisted of 1,249,737 men and 1,271,385 women

iliary OLS regressions have very high R^2's suggesting that substituting previous income by current income is acceptable.[21] The 14 'cumulative' logit models show that older, poorer and hospitalized individuals have a significantly lower probability to survive the next year. This is in line with *a priori* expectations, but not very informative to understand the process of selective mortality. For the latter, the interactions between age and income, and between income and hospitalization are more revealing. Our estimates suggest that high income has a stronger positive survival effect for older individuals. We also find that a hospitalization particularly reduces survival chances of individuals in the poorest income group. This is confirmed by Figure 2.2 which visualizes the magnitude of these interaction effects. It plots the relationship between age and the IPW's for the year 1999

21. For the individuals belonging to both the RIS and SLC, we also compared the IPWs resulting from the 7 'cumulative' logits with those obtained from the auxiliary regressions. We found that the IPWs were almost identical and revealed similar patterns in terms of interactions between income and age and income and hospitalizations.

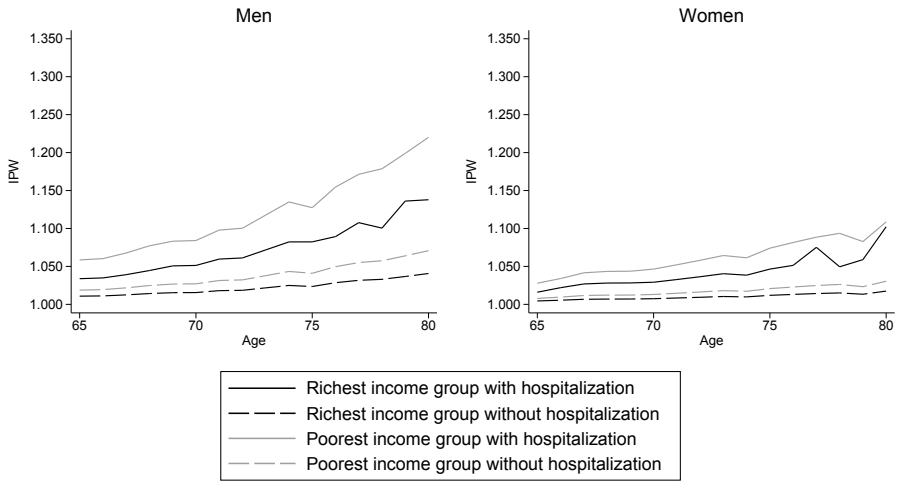


FIGURE 2.2: Average IPWs due to selective mortality between 1998 and 1999 for men and women by income and hospitalization

for 4 groups, i.e. individuals belonging to the highest (€30.000 and higher) and lowest income group (€0-€12.500), interacted with having been hospitalized or not. We show this separately for men and women. The larger impact of being hospitalized on dropout for those in the lowest income group is clear from the larger difference in IPW's between those hospitalized or not. For 70 year old males, for example, the one-year survival difference between the hospitalized and non-hospitalized is 50% larger in the lowest than in the highest income group. The patterns in Figure 2.2 show, however, that the selective mortality weights are not large, certainly not for those under age 70, and their application may have a limited effect on the ill-health life cycle profiles displayed in Figure 2.1. We will turn to this in section 2.4(d).

(c) Calculating IPWs – selective institutionalization

Construction of IPWs for selective institutionalization and correction for its effect is done separately to examine whether attrition effects of mortality and institutionalization differ by age, income and prior hospitalization. The SLC is a representative sample of the Dutch population living in private households only[22], but we can identify SLC members that moved to an institution by exploiting the linkage between the RIS and the municipal population register (MPR) which includes information on the residential status of all individuals living in the Netherlands.

22. As in the case of selective mortality and – as far as we know – in all other literature on this topic, this implies we cannot analyze pre-SLC selective institutionalization.

The IPWs for selective institutionalization are constructed in the same manner as those for selective mortality, i.e. by first running 'cumulative' logit models in the RIS[23] and next running auxiliary OLS regressions to link to IPWs obtained in the RIS to the SLC. The MPR defines an 'institutional household' as a nursing home, a revalidation center, a psychiatric hospital, a long term care institution for the mentally ill or a prison.[24] We use whether the individual lives in a private household in year *t* as the dependent variable in the logit models, and employ the same set of explanatory variables as in the models for selective mortality: one year age dummies, 4 absolute income groups, hospitalization in the previous year and the same one-way interactions. Since the dependent variable is not defined for individuals that die during 1998-2005, we remove all these individuals from the analysis that corrects for selective institutionalization, i.e. both for the IPWs and the ill-health life cycle profiles.[25]

Table 2.2 shows that 0.71 percent of males and 1.51 percent of females of the 1998 RIS sample have dropped out due to institutionalization by 2005. The logistic regressions from which the IPWs are derived show smaller effect sizes of income, age and hospitalization on institutionalization (compared to the effects on mortality) and indicate an insignificant interaction between income and hospitalization. For reasons of space we do not show the regression results, but we plot the IPWs for the year 1999 against age for the lowest (€0-€12.500) and highest income group (€30.000 and higher), interacted with hospitalization (similar to Figure 2.2). Figure 2.3 shows smaller IPW values compared to those of mortality in Figure 2.2, and smaller absolute differences between the lines defined by income and hospitalization. Nevertheless, we still observe evidence of selective institutionalization since the effect of having been hospitalized seems to matter mostly for the lowest income group, although the interaction between income and hospitalization was insignificant.

23. Unlike mortality, institutionalization is not always an absorbing state. In practice, re-entry into a private household is very rare (e.g. former prisoners). Therefore, we only considered individuals 'institutionalized' when they remained in an institutionalized household until 2005. Individuals moving back to a private household for only one year before moving to an institution for multiple years afterwards are assumed to be in an institution for the whole period.

24. The MPR does not separately identify each type of 'institutional household'.

25. We also constructed the IPWs for the combination of selective mortality and institutionalization to allow for potential dependencies between mortality and institutionalization, but this did not change the qualitative conclusions reached in this paper.

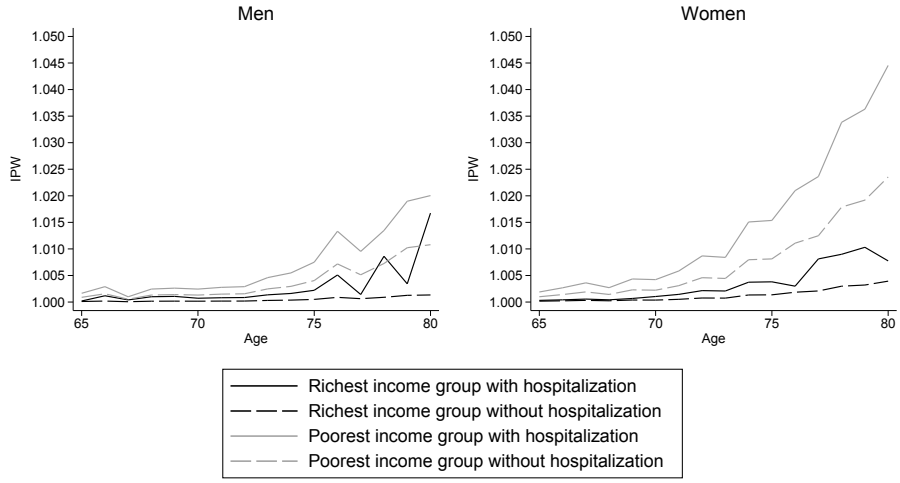


FIGURE 2.3: Average IPWs due to selective institutionalization between 1998 and 1999 for men and women by income and hospitalization

(d) The life cycle profile of health across income groups after correction for selective mortality and institutionalization

In this section we examine whether the ill-health life cycle profiles reported in Figure 2.1 change after correcting for selective mortality and institutionalization. We adopt the procedure discussed in section 2.4(a) and apply the IPWs of sections 2.4(b) and 2.4(c).

Given our findings in section 2.4(b) and 2.4(c), we would expect the ill-health life cycle profile of the poorest quartile to shift upward after correcting for selective dropout, and that of the richest quartile to remain the same (or slightly shift down). This would imply that the convergence of the ill-health levels of the rich and poor – as evident in Figure 2.1 and in line with the age-as-leveler hypothesis – would turn out weaker. For selective mortality, we also expect (a) the effect to be larger at higher ages –especially for the lowest quartile – because of the interaction between age and income; and (b) the effect to be larger for the poorest quartile since this group has poorer health (as proxied by hospitalization) which leads (for the same income) to a lower survival probability. Given the lower absolute values of the IPWs for institutionalization and the insignificance of the interaction between income and hospitalization, we expect a more limited correction of the ill health life cycle profiles.

Figure 2.4 panels a (males) and b (females) reproduce the ill-health life cycle profiles reported in Figure 2.1 and compare these with the profiles resulting after correction for selective mortality. Remember that the IPW corrections matter for the construction of both the income quartiles and the average ill-health levels per income quartile. For males, we find an effect in the expected direction from age 50 onward for the lowest income

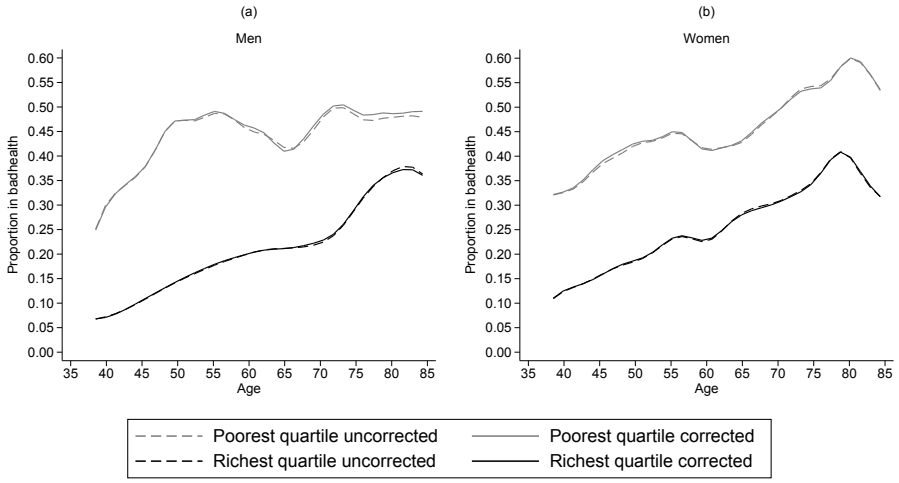


FIGURE 2.4: Mortality-corrected patterns of poor health for men and women over the life cycle in the Netherlands (1998-2005)

quartile, except around the age of 63. We hardly see any effect for the richest income quartile. For men the maximal effect is reached around the age of 80 and increases the gap between the poorest and richest group by around 28 percent, for women this is at age 86 when the gap increases by 8%. While this is a substantial effect, it seems unrelated to the health improvement observed between ages 55 and 65. We thus conclude that selective mortality seems to somewhat reinforce the convergence of income-differences in ill-health at older ages, but that it cannot explain why the poorest income quartile experiences an increase in their health around age 55-65. We also observe a downward shift for rich males after 80, but because of smaller cell sizes we must be cautious with interpreting this as a fact. For females, the effects are more limited and basically show that selective mortality is unimportant to explain income differences in the evolution of female health. Figure 2.5 panels a (males) and b (females) show the corresponding results for selective institutionalization. The effects are generally much smaller and tell us that selective institutionalization is an unimportant phenomenon which is most likely explained by the low number of institutionalizations compared to the number of deaths (see Table 2.2) and the universal coverage for long term care in the Netherlands. A small effect is visible around the age of 75-80 for women in the richest quartile, who seem to improve health after correction. This decrease is about 6% and occurs at the age at which many Dutch women move into nursing homes. However, this effect is not visible at older ages, which could indicate that women quickly die after entering a nursing home between 75 and 80 or just be a result of the small sample size at older ages.

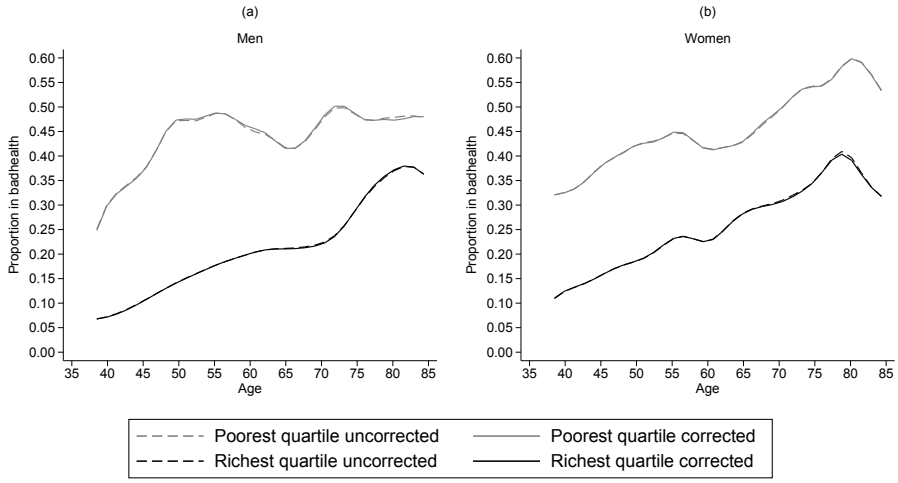


FIGURE 2.5: Institutionalization-corrected patterns of poor health for men and women over the life cycle in the Netherlands (1998-2005)

2.5. DISCUSSION AND CONCLUSION

This paper focuses on one possible explanation for a well-documented but not well understood pattern in the life cycle profiles of health by SES: differences in health across SES groups tend to diverge from young adulthood onwards and widen until late middle age, after which these health differences start converging again. We estimate the extent to which selective mortality and selective institutionalization explain the convergence of health trajectories by SES at higher ages.

The issue has been investigated by others, but we adopt a reweighting approach to correct for selection effects that overcomes some of the earlier deficiencies. We use a series of repeated cross-section surveys covering the period 1998-2005 to construct longitudinal data by birth year/gender/income groups to obtain life cycle profiles of health for different income groups. Linked administrative data are used to estimate individual probabilities of mortality and institutionalization in the seven year period for more than 2.5 million individuals to obtain precise estimates. Any selection biases due to mortality and institutionalization in the life cycle profiles by income are then corrected using inverse probability (re)weighting. The combination of four essential features distinguishes our approach from earlier ones: (a) the IPWs depend on the experience of both survivors/non-attritors and decedents/attritors; (b) we allow attrition to depend on the interaction between prior SES and prior health status; (c) we allow for dynamics in (the effect of) the explanatory variables underlying the IPWs which makes the assumption of 'missing at random' more plausible; (d) we do not only study the possible confounding effects of selective mortality but also of selective institutionalization.

Our findings are as follows. First of all, we find evidence of both selective mortality and institutionalization at higher ages. Attrition is not random: older, poorer and unhealthier (proxied by hospitalized) individuals are significantly less likely to survive the next year and more likely to be admitted to a long term care institution. While all of these selection effects are significant, they are not very large. For example, the probability of dying in the first year is at the oldest ages 4 percentage points higher for the poor than the rich, but this difference is raised by less than 2 percentage points when poor health is accounted for. These are, however, still relatively small effects, given that the probability of dying at the oldest ages within a year after a hospitalization is higher than 20%.

Secondly, and more importantly, we allow the selectivity to be not only health- and age-related but differentially so by income groups. We do this by including interactions between lagged health, age and income and find that high income has a stronger positive survival effect for older individuals. We also find that lagged health (proxied by hospitalization) decreases survival chances most for the poorest income group. While these effects are relatively large, their absolute impact on survival is limited. If sufficiently large, this heterogeneity in effects may help explain the observed lifetime patterns.

Thirdly, the corrections for selective mortality and institutionalization adjust the life cycle patterns of health by age and income in the expected direction, and increase the gap between rich and poor with maximally 28 percent. This reduces the convergence of the 'uncorrected' ill-health life cycle profiles between poor and rich. Among males, the effect is visible from the age of 50 onwards, but most pronounced around the age of 80. The same is generally true for females but the magnitude of the effects are much smaller and only emerge above age 75. Corrections for selective institutionalization are smaller at all ages mainly due to the lower average rates of institutionalization compared to the average death rates and the lower magnitude of income effects.

All in all, these findings imply that – despite the dropout being selective by both health and income – the healthy survivor effect cannot be the main explanation of the first diverging and then converging pattern in the low-to-high income difference in health by age. Other mechanisms must be responsible for its occurrence. One possibility is reverse causality: rather than the health of those with high incomes deteriorating faster with age, it could be that lower health reduces income through individuals dropping out or reducing their participation in the labor force. Our data do not permit a proper test of this hypothesis but recent work by García-Gómez et al. (2013) on the Netherlands shows that an acute hospital admission lowers employment and personal income without subsequent recovery, and the impacts are stronger at the bottom end of the income distribution. Their findings suggest that reverse causality – from health to income – is another important mechanism driving the income-health association and which is worthwhile exploring to further unravel the mechanisms behind the hump shaped pattern of the life cycle pattern of the SES health gradient.

Chapter 3

The SES-BMI gradient in China: Rotating with age

ABSTRACT

Background:

Levels of body mass index (BMI) are on the rise in China and are related to socio-economic status SES. The SES-BMI gradient is already present in childhood and carries over to later life inequalities.

Aim:

To investigate the linkage between maternal educational attainment and adulthood BMI and reveal alternative potential pathways.

Data:

The longitudinal data come from the China Health and Nutrition Survey (CHNS) which provides height and weight measurement from 1989 to 2009 and includes household data on maternal education.

Methods:

We use ordinary least squares, linear probability and quantile models to investigate the determinants of BMI, overweight, obesity and the BMI distribution. We perform stepwise analyses including possible intermediating factors between childhood SES and adulthood BMI.

Results:

We find evidence for an effect of maternal education on BMI levels later in life. However, the impact rotates with age, showing that BMI levels are lower for high SES individuals at young ages, but higher at older ages. A large part of the association seems to run through individuals' own educational attainment, which also displays a rotating relationship with BMI with age. Results are similar for overweight, but not for obesity, for which no relationship with SES in either child- or adulthood is found. The results of our quantile regressions suggest that SES is important across most of the BMI distribution, but not at its tails.

Conclusion:

Early life conditions in China co-determine adult BMI levels. The rotation of the gradient highlights the complexity of this relationship and the need for further research into the social determinants of health in China. Explanations might be sought in changes in education itself over time, diminishing effects or differential age effects for other determinants like gender, region and work status.

3.1. INTRODUCTION

With the good often also comes the bad, and China is no exception. China's development in the last three decades has brought great improvements to many Chinese households, but has also increased the incidence and prevalence of diseases of affluence. Specifically, overweight and obesity levels keep rising and have become a major threat to Chinese public health (Wang et al., 2006b). High body mass index (BMI)[26] has been linked to many diseases in the field of cardiology, endocrinology, neurology and others. Until now, the impact in China has been relatively low, certainly in comparison to the United States, but projections suggest that in 2030 more Chinese will be overweight than people in all established market economies together (Kelly et al., 2008).

Only BMI increases that have negative health outcomes are problematic. For many individuals with healthy levels of BMI this is not the case (Deaton and Paxson, 1998a). This complicates the analysis of the distribution of BMI, because not all differences in BMI are necessarily unwanted. This is especially true in China where the distribution has been shifting very rapidly. In 1989 only 9.1% of the Chinese were overweight and a similar proportion (8.7%) was underweight. By 2000, the share of overweight individuals had doubled (19.7%) and the share of underweight Chinese had decreased to 6.5% (Wang et al., 2006a). Nonetheless, there is sufficient evidence to be concerned about increasing BMI levels in China at all ages and across the entire BMI distribution (Wang et al., 2006a). If this trend continues, those currently within the healthy bounds of BMI might be overweight or obese in the near future. To adequately identify the groups who are most at risk it is important to identify the drivers of the current growth in BMI.

In Western countries, one key correlate of high body weight is socio-economic status (SES) (McLaren, 2007). Generally, the so called SES-BMI gradient links low SES to high BMI, overweight and obesity prevalence and consequently socio-economic inequalities in health. The gradient is less clear-cut in China with respect to its direction (Kim et al., 2004) and regional variation (Chen and Meltzer, 2008; Van de Poel et al., 2009). This paper examines how adulthood BMI is related to childhood SES and investigates the pathways through which this relationship might operate. Additionally, we examine the heterogeneity of the SES-BMI relationship at different levels of BMI by looking at average levels of BMI, prevalence of overweight and obesity as well as the entire distribution of BMI.

26. BMI is the most commonly used measure for unhealthy weight levels and is calculated as weight in kilograms divided by height in meters squared. Although being the most common measure BMI is heavily criticized for not distinguishing between fat-free body mass and fat body mass (Burkhauser and Cawley, 2008). In this paper, however, we only consider BMI as it is still the number one measure used and it makes our results comparable to that of others.

A better understanding of the dynamics between socio-economic status and weight in China is important to anticipate in what direction China may be heading. A key issue when unveiling these dynamics is the direction of causality. High SES might lead to lower levels of body weight, but low levels of body weight could also be the reason that people are able to attain higher SES. Another possibility is that third factors, observed or unobserved, may exist that affect both SES and BMI. For a proper understanding of what drives the increases in body weight these conflicting effects need to be disentangled. One possible solution is to focus on SES during childhood and BMI during adulthood. The former is primarily determined by parental characteristics and surely not by the latter. A relationship between the two is not necessarily causal, but childhood circumstances have been shown to have long lasting effects on future socio-economic circumstances and health (Case et al., 2005). The direct relationship between childhood SES and adulthood BMI does not show the dynamics behind rising BMI levels. To uncover these, the focus should also lie on what happens between childhood and adulthood. The relationship between childhood SES and adulthood BMI could still be determined by third factors such as genes and early life living circumstances. Other third factors may be less relevant now. For example, later life living circumstances cannot affect SES of the parent, but only the BMI of the child.

We explore the relationship between childhood SES and adulthood BMI and model possible pathways through which the association might run. We find that a relationship exists, albeit an ambiguous one. The relationship is not found to be linear, and to vary with age. This SES-BMI gradient rotates and its effect depends on whether early adulthood health and adulthood SES are controlled for. The next section discusses how the increases in body weight in China might be linked to changes in SES status and what potential mechanisms may mediate the childhood SES - adulthood BMI gradient. The third section describes the data and our analysis. Section four presents the results and the last section concludes.

3.2. SOCIO-ECONOMIC STATUS AND BMI

The pace at which high body mass is becoming a problem is remarkable. Only half a century ago China faced the worst famine in modern history. In spite of the upward trend in average BMI levels, undernutrition is currently still a concern, which is reflected in a "double burden of malnutrition". Consequently, the SES-BMI gradient in China is less unambiguous than in the west where a clear disadvantage exists for lower SES groups. In China, however, only those with high SES have been able to overconsume for a longer period and only very recently this has become a potential problem for most Chinese (Shankar, 2009). This shift from a concentration of overweight and obesity among the richer and better educated to those who are worse-off is not uncommon in developing countries in transition (Jones-Smith et al., 2011).

Philipson and Posner (2003) provide a rational choice model to link increases in body weight to technological changes in a society, which is helpful to understand China's rapid increase in BMI levels. The model explains that the utility derived from body weight is inversely u-shaped. As a result, rational individuals select a weight that makes him or her happiest given the costs and benefits of calorie intake and consumption. Since food is relatively expensive in underdeveloped countries, like China a few decades ago, additional income will lead to higher food consumption and thus increased body weight. Yet, in more developed countries, like present-day China, weight levels have risen considerably as a result of higher incomes levels and lower relative prices of food. Individuals will then seek to reduce their weight by limiting their food intake and increased exercise to compensate for the reduced physical activity that follows from the sedentary nature of most jobs in developed countries. Clearly, these stimuli operate on the individual level and the role of income will be different in areas with different average levels of development. This is illustrated by the fact that in China there are still many who believe that excess body fat is healthy and a signal of wealth (Wu, 2006). This process has resulted in the concentration of high levels of overweight among those with high incomes in economically less developed regions and higher concentrations among poor individuals in more developed regions (Tafreschi, 2011).

The increased BMI levels in China have been linked both to higher energy intake and lower levels of physical activity. Energy intake has risen partly because of the higher energy density of food (Kant and Graubard, 2005) and the importance of edible oils in modern diets (Ng et al., 2008). These types of food are more commonly available nowadays because of price drops as also explained by Philipson and Posner (Lu and Goldman, 2010). Additionally, recent studies show a drop in average level of weekly physical activity in China by 32% (Ng et al., 2009). Evidence suggests that - in contrast to the United States - this might be a more important driver of the increases in body weight than energy intake (Bell et al., 2001; Cutler et al., 2003). More specifically, this discrepancy with the United States seems to be related to the role of sedentary jobs, which are most common among high SES individuals and seem to be an important contributing factor to BMI increases in China, but not in the United States (Gutierrez-Fisac et al., 2002; Kim et al., 2004; Monda et al., 2007).

Ultimately, all variation in body mass is explained by the amount of energy intake and energy consumption. Cutler et al. (2003) estimate that an imbalance of only 100 to 150 calories a day could account for the entire increase of body weight in the U.S. between 1980 and 2000 (Cutler et al., 2003). For China the current gap between energy intake and consumption is 45 calories (Zhai et al., 2008). So, theoretically, modeling these two variables would suffice to explain the increases in BMI in China. However, it is difficult to measure both variables accurately and more importantly, it would not reveal the dynamics between SES and BMI. Different pathways in which SES is associated with overweight

and obesity have been proposed. For example, it has been suggested that higher levels of education are associated with better knowledge about good nutrition and consequently superior body mass control. Conversely, those with high levels of body mass could face a wage penalty compared to those who a healthier body weight. Another explanation for an association between overweight and high SES could be that only those with high levels of income would be able to afford the food intake necessary to develop overweight. Especially, in a developing economy like that of China this could have been an important pathway in the past decades.

Summarizing, China is currently different from Western countries, but the changes in BMI levels follow a process that is not uncommon for developing countries. These changes are the result of both energy intake and energy expenditure and could very well be related to socio-economic status.

3.3. ANALYSIS

(a) Pathways from childhood SES to adulthood BMI

We assess the relationship between childhood socio-economic status and BMI during adulthood. Like others we use maternal education as an indicator of childhood SES, which often correlates strongly with child health (Currie et al., 2007). We summarize the dynamics underlying the effects of maternal education on BMI in Figure 3.1. The pathways are not mutually exclusive and may run through either childhood health or adulthood socio-economics status and all run through life styles. In our analyses we will only use adulthood BMI based dependent variables. Consequently, we only measure those relationships that end at the bottom of Figure 3.1. Besides, we will not be able to measure lifestyles so accurately to explain all variation in BMI. Therefore, dashed lines have been added to represent the effects we do estimate. No examples are possible for these pathways, since all variation in BMI would depend on the energy intake and expenditure as measured by the lifestyles, at least theoretically. Models that include lifestyles and yet show a significant impact of one of these dashed lines, illustrate the limitations of the measurement of energy intake and expenditure, rather than true effects. These dashed lines do, however, reveal the impact of the different pathways on the relationship between childhood SES and adulthood BMI.

Maternal education may influence future BMI in a number of ways. First of all, maternal education can directly affect BMI if, for example, higher educated mothers teach children

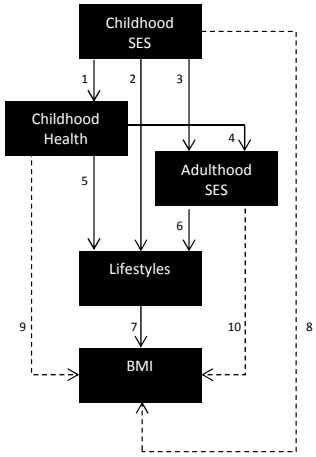


FIGURE 3.1: Possible pathways from childhood SES to adulthood BMI

healthier ways of living, which are used far into adulthood[27]. This is represented by the lines 2 and 7 in Figure 3.1. If childhood SES affects health already during childhood, this effect can carry over to adulthood (e.g. Currie, 2009). It could limit someone's ability to exercise sufficiently and sustain a healthy body weight. Poor childhood health could have also been reflected in stunting, which would affect height – the denominator in the calculation of BMI - for the remainder of the lifespan. These effects are described by lines 1, 5 and 7 in Figure 3.1. Adulthood SES may function as an additional mechanism, because it is associated with both childhood SES and BMI. Higher educated individuals have lower discount rates, which may translate into healthier life styles now and better health in the future, i.e. lines 3, 6 and 7. They are also more likely to be able to afford good nutrition, see Philipson and Posner (2003). Adulthood SES can also result from bad childhood health, i.e. the pathways described by lines 1, 4, 6 and 7. Physical activity may also be important here because of the sedentariness of jobs for the higher educated, but this effect is captured in the lifestyles.

We analyze each of these pathways, but measure the direct effect on BMI (dashed lines) rather than through life styles (solid lines). As mentioned before, BMI is the result of a combination of energy intake and expenditure reflected in these lifestyles. However, we are not necessarily interested in the effects of childhood SES on the amount of food

27. Please note that in this paper we call all effects of maternal education as effects of socio-economic status during childhood. We ignore that the impact that maternal education is not always the result of the related socio-economic status. The effect described here, for example, is unrelated to the economic resources of the family during childhood. One could interpret this as a social effect of the maternal education. Through the remainder of this paper, we do not distinguish between effects of maternal education that can be linked to economic resources and those that cannot.

consumed and the impact of physical activity. We are simply interested in the net impact on BMI, because BMI is associated with negative health outcomes. Furthermore, the relationship between lifestyles and BMI is complicated and includes not only the calories consumed and the energy used in physical activity, but also energy consumed through maintaining the body, i.e. basal metabolism, and processing food consumed, which combined account for 70% of energy consumption (Cutler et al., 2003). These two types of energy expenditure are difficult to measure and we therefore opt to, first, measure the effects on BMI without considering lifestyles.

(b) Models to explore pathways to high BMI

We explore these pathways empirically by running consecutive ordinary least squares (OLS) regression models on BMI. Although the CHNS is a panel dataset, we do not use panel models to estimate these relationships since our primary interest is in the effect of the time invariant variable maternal education. In a fixed effects model, its coefficient would be completely absorbed in the individual fixed effect. Therefore, we opted to present pooled OLS models.[28] A first set of models explains BMI using only a set of background demographics. These include dummies for the interaction between gender and marital status. Marital status has been reported to have both positive and negative effects on levels of BMI in western settings (Sobal et al., 2009) and is suggested to raise levels of BMI in China (Hu et al., 2002). Furthermore we used age linearly and squared in the models, as well as time and region dummies as background characteristics[29].

Next, we add maternal education to assess the overall relationship between childhood SES and adulthood BMI, i.e. pathway 8 in Figure 3.1. This pathway is central in our analysis, because we assess how it changes when we account for other pathways. First we account for adulthood SES to assess pathway 10 and its impact on pathway 8. Then, we add caloric intake and expenditure as life style indicators, to examine the importance of pathway 8 after including pathway 7 in addition to pathway 10. We also look at the impact of health at younger ages on BMI later in life, i.e. pathway 9, but do so using a limited sample because of data restrictions (see below). Note that we do not evaluate pathways 1 through 6 which, among other things, include the impact of childhood SES on health

28. We did also run random effects models to see how much our results would be affected. In most models few differences were observed. Only in the more elaborate models the size of certain coefficients changed, but generally not the sign or level of significance.

29. Cohort effects could also explain BMI trends. These could result from younger individuals changing their diet and activity pattern faster than elderly may also be important or long lasting effects of historic events, such as the great famine (Luo et al., 2006), the Sino-Japanese war, the Cultural Revolution or the start of the economic reforms. We excluded these cohort effects because of the impossibility to simultaneously correct for all three of these effects (i.e. age, time and cohort) in regression models later on. Since, we are not interested in the magnitude of these effects, but only correct for them, we leave out cohort effects to be consistent.

earlier in life and adulthood SES. These effects are implicitly included in pathways 7, 9 and 10.

While trends in *average* BMI levels are important, trends in the prevalence of overweight and obesity are even more important for health matters. A SES-BMI gradient around the thresholds for overweight and obesity will have a larger impact on SES-health disparities than a SES BMI gradient at lower BMI levels. In case the trend of increased average BMI levels in China ends, this will in the end be all that matters in this context. We therefore repeat the analysis for the prevalence of both overweight and obesity. We use linear probability models for ease of interpretation of the coefficients. Probit models, delivered qualitatively similar results.

These models assume that a large part of the variation in BMI is irrelevant, since they treat having a BMI just above the respective threshold the same as one that is way above the threshold. Health risks are clearly larger in the latter case, which may lead to over- or underestimation of the health impact of the SES-BMI gradient. We could have opted to look at z-scores, which show by how much healthy weight levels are exceed for certain (SES) groups (World Health Organization, 1995). These would, however, again assume that effects are constant over all z-scores. We therefore use quantile regression to estimate differential SES effects along the entire BMI distribution below and above the thresholds of overweight and obesity (Sturm and Datar, 2005). These quantile regressions will provide additional information on how SES affects the entire BMI distribution. They do not provide marginal effects on the individual, but on the quantile. The coefficients for, e.g. the 5th, centile will indicate how much heavier the 5th centile would be if maternal education had been higher in the entire sample. This does not mean that those individuals currently in the 5th centile would still be the ones in that centile and thus have a higher BMI. In other words, the quantile regression provides the effect of a set of determinants on the BMI of the conditional quantile rather than people in the quantile themselves.

3.4. DATA

We use the *China Health and Nutrition Survey* (CHNS) to evaluate possible pathways of maternal education effects on BMI in adulthood. The CHNS is a panel study that started in 1989 and is still ongoing. The latest, i.e. eighth, available wave was collected in 2009. In addition to good measurements of weight and height, the CHNS contains a broad set of questions on social circumstances, health levels and behavior and diet. Our sample is limited to adults over the age of 18, although we do collect information from their childhood through data collected in waves before their 18th birthday.

(a) Measurements of body mass

Throughout the analysis we use body mass index to evaluate individual weight levels. In each wave of the CHNS, both weight and height[30] were not self-reported, but measured by the interviewers. This should lead to more accurate calculations of BMI, especially at older ages (Kuczmarski et al., 2001). Commonly, BMI thresholds of 25 kg/m^2 and 30 kg/m^2 are used to classify overweight and obesity. However, in China and other Asian countries there is less consensus about the thresholds of overweight and obesity, because Chinese have higher levels of body fat compared to Caucasians at similar levels of BMI and the thresholds applied in the west are therefore not suitable (Deurenberg-Yap and Deurenberg, 2003). Even between studies conducted among Asians differences exists in the selected cut-points. In Asia BMI levels of 23 kg/m^2 already lead to increased risks of type 2 diabetes and cardiovascular diseases, at 27.5 kg/m^2 and above Asians are already at a high risk for developing these diseases (Barba et al., 2004). Therefore, we use these lower cut-off values for overweight and obesity.

(b) Maternal education

Maternal education is only limitedly available in the CHNS, because maternal education is only attainable through direct questions to the mother. In other words, the mother also must have participated in at least one wave. In the CHNS households are sampled so mainly children and adolescents will be observed along with their mothers. The long follow-up in the CHNS allows us to observe sufficient respondents both during child- and adulthood. Maternal educational attainment was assessed at the end of childhood (i.e. 18 years) or as close as possible. This is an assumption that increased maternal education during childhood may be an indicator of the childhood SES and will contribute to BMI levels later in life. It is however not certain that changes in SES at the end of childhood are as important as the SES at the youngest ages, which would make maternal educational attainment at birth a better indicator. The follow up in the CHNS, albeit a relatively long one, is not quite long enough to observe children and their mother at age 0 and measure their BMI during adulthood. What is more in over 85% of the respondents their mother has the same education at the first wave as at the last wave, so this choice will most likely only have a limited effect on the outcomes. Table 3.1 shows that half of the respondents in the CHNS have missing information on maternal education and over a third is excluded because of other reasons, such as missing BMI (22%) and being under 18 (10%)

30. Height should be near constant within our sample of Chinese adults, but some inconsistencies exists within the CHNS. Of our panel about 75% of height observations is equal to an individual's modal height. Of those 25% with variation in their height level, 50% of the observations only diverge 1.5 centimeter from their modal height.

TABLE 3.1: Dropout due to missing information on different variables by wave

		Dropout					
Wave	Total	No maternal education	No height or weight	No education	No information on lifestyles	Younger than 18	Included
1989	15,891	8,332	4,885	1,496	66	260	852
1991	14,757	7,603	1,590	1,232	110	2,568	1,654
1993	13,868	7,146	1,566	1,055	69	2,378	1,654
1997	15,779	8,395	2,765	788	141	2,245	1,445
2000	16,983	9,254	3,180	1,172	148	1,753	1,476
2004	16,095	9,042	3,492	425	85	1,456	1,595
2006	18,687	9,682	5,843	389	81	1,147	1,545
2009	20,828	11,002	6,464	425	62	1,031	1,844
Total	132,888	70,456	29,785	6,982	762	12,838	12,065

We have asserted that the difference between the samples with and without maternal education will not lead to different conclusions about the SES-BMI gradient. The BMI distribution conditional on age and gender is similar for those with known and unknown maternal education. The average age however, is lower because older individuals are less likely to live with their mothers. The share of females is also smaller than that in the full CHNS, which may be the results of elder mothers to live in a household with a son rather than a daughter.

(c) Measures of contemporary SES

We use respondents' own educational attainment as a measure of adulthood SES. Education has been reported to be the best measure of SES in this context (Zhang and Wang, 2004). It is associated with lower discount rates as those who are willing to invest in their education are also willing to invest in their future health through better life styles (Borghans and Golsteyn, 2006)[31].

(d) Region

We use dummy variables of region to correct for differences in levels of development in China. In China, like in many other settings, urban areas have developed faster than rural areas. In addition, traditionally Chinese coastal provinces have been more developed than inland provinces. In the initial stages of China's economic reforms these coastal provinces also received a preferential treatment over the other provinces (Chen and Fleisher, 1996). Since anticipated spillover effects did not happen, inland provinces still face a large deficit in development compared to other regions. We therefore included

31. When SES was measured using income, a weaker correlation with BMI was found and our main findings with regard to the relationship between maternal education and adulthood SES remained unaffected.

dummies for the interactions between urban/rural and coastal/inland regions to account for geographical differences in average development.

(e) Lifestyle indicators

Lifestyles in the context of BMI are choices that influence one's body weight. Such indicators would ideally measure any choice by an individual that increases or decreases his or her energy intake and expenditure. Clearly, not all such choices are available in a survey. Therefore, the most important choices with respect to intake and expenditure are chosen. For intake the calories consumed in one's diet will mostly determine the level of energy. For expenditure, a person's daytime activities, i.e. type of work, are the most important choice. Details on how nutritional information was collected in the CHNS can be found in Cui and Dibley (2012). We are only interested in the energy component of diets and therefore only focus on caloric consumption[32]. Since the required caloric intake differs between males and females, we added interactions with caloric intake[33]. We also constructed a dummy variable to account for energy expenditure based on individuals' main daily activity. Activity level was assessed by CHNS interviewers and categorize in five levels: very light (working in a sitting position); light (working in standing position); moderate (e.g. student or metal worker); heavy (e.g. farmer) or very heavy (e.g. logger or miner). Our dummy for low activity includes those in the light and very light category.

(f) Childhood and early adulthood health

The follow up of the CHNS is not long enough to observe all respondents both during childhood and adulthood, which complicates measurement of health during childhood. As a result we can only approximate childhood health by early adulthood health. Besides, we still can only provide the analysis of pathways associated with early adulthood health as a side analysis, because early adulthood is still only available for a limited number of respondents. We use a self-rated health (SRH) measure, which was included form 1991 to 2006. The SRH was phrased as follows: "Right now, how would you describe your health compared to that of other people of your age?": Excellent, Good, Fair or Poor?". To include all health outcomes during early adulthood and maximize the number of observations we used SRH measures at the end of childhood, i.e. age 18. Because we do not observe all respondents at age 18, we include responses to the SRH question as close as possible up to the age of 25. Clearly outcomes at 25 are no longer only the result of events

32. We also considered whether it was of importance from which nutrients these calories were derived, but found no support for a more comprehensive role of diet in our model.

33. More comprehensive estimates to calculate total daily energy expenditure (e.g. Willcox et al., 2007) based on the basal metabolic rate and the Harris-Benedict equation (1919), were not feasible, since these rely on individual weight and height, which are heavily endogenous.

during childhood, but a stricter inclusion criterion would limit our sample size too much. We only have information on SRH before the age of 25 for 5,478respondents. This sample is too small to properly evaluate all relevant pathways, but might give an indication to whether health outcomes during childhood are important determinants of BMI levels later in life.

3.5. RESULTS

We first look at the composition of the sample of included respondents classified by their mother's educational attainment and their own. Table 3.2 shows that the majority (53%) of mothers has received no formal education and less than 20% went to school after finishing primary school. BMI levels of the first group are higher than those of groups with higher educated mothers. These differences, however, seem small given the associated standard deviations. Children of higher educated mothers are younger and have attained higher levels of education themselves. Higher maternal education also relates to higher adulthood SES in terms of income. Equivalent household income for children of mothers without education is ¥ 9,157, while this is more than double (¥ 19,961) for children of mothers with 13 or more years of education. Furthermore, most children from uneducated mothers live in inland rural areas, whereas maternal education is highest in urban areas. The bottom panel suggests that no relationship between respondents' own education and BMI exists. Other associations seem similar to those observed for maternal education. Higher educated respondents are younger, richer and more likely to live in urban areas.

We investigate the age effects in relation to maternal education in Figure 3.2[34]. The figure shows two things: (i) the age gradient is much stronger than the SES gradient and (ii) the gradient is not the same at all ages; differences in BMI between age groups are different at different levels of BMI. This can be interpreted as a rotation of the gradient as a result the SES-effect is different by age; at young ages *low* SES is associated with higher levels of BMI, whereas at older ages *high* SES implies higher levels of BMI. This is somewhat in line with the findings of Case et al. (2005) of a steepened gradient at older ages, but the gradient in Figure 3.2 rotates in the opposite direction; from advantageous BMI levels for high SES at young ages to higher levels at older ages. The rotation implies that, over the life cycle, events related to childhood SES still occur that affect BMI. This is because after childhood, the BMI still changes at a different rate for those with high and low SES during childhood.

34. We only include these three types of maternal schooling in this figure as they include 94% of all respondents; the 6% of the highest educated mothers are included in the subsequent analyses.

TABLE 3.2: Descriptive statistics by level of maternal education and respondent's own education

	Education		N	BMI	Age	Female	Income	Educ. years resp.	Educ. years resp. mother	Inland rural	Inland urban	Coastal rural	Coastal urban
Maternal education	None	mean	6,438	22.0	35.0	0.26	9,157	8.0	0.0	0.51	0.21	0.20	0.08
	(0 years)	% / std. dev.	53%	3.0	12.8	0.44	11,451	3.3	0.0	0.50	0.41	0.40	0.27
	Primary	mean	3,579	21.7	27.5	0.33	11,804	9.6	4.3	0.46	0.25	0.18	0.11
	(1-6 years)	% / std. dev.	30%	2.9	8.6	0.47	13,148	2.6	1.6	0.50	0.43	0.38	0.31
	Lower middle school	mean	1,328	21.6	25.4	0.34	15,217	10.8	8.6	0.37	0.26	0.19	0.18
	(7-9 years)	% / std. dev.	11%	3.2	7.6	0.48	19,370	2.6	0.7	0.48	0.44	0.39	0.39
	Upper middle school	mean	633	21.6	25.5	0.38	18,293	12.3	11.5	0.24	0.41	0.12	0.23
	(10-12 years)	% / std. dev.	5%	3.3	7.7	0.48	14,614	2.5	0.6	0.43	0.49	0.33	0.42
	College	mean	87	21.3	26.0	0.37	19,961	14.0	15.2	0.02	0.30	0.00	0.68
	(13+ years)	% / std. dev.	1%	2.6	6.6	0.49	15,051	2.1	1.0	0.15	0.46	0.00	0.47
Own education	None	mean	408	22.0	44.3	0.46	5,997	0.0	0.5	0.48	0.19	0.26	0.07
	(0 years)	% / std. dev.	3%	3.1	16.9	0.50	8,247	0.0	1.6	0.50	0.40	0.44	0.25
	Primary	mean	2,309	21.9	35.3	0.30	7,901	5.1	1.1	0.61	0.17	0.18	0.04
	(1-6 years)	% / std. dev.	19%	2.9	13.9	0.46	12,519	1.5	2.4	0.49	0.37	0.38	0.20
	Lower middle school	mean	5,437	21.8	29.5	0.27	10,200	8.9	2.5	0.49	0.21	0.21	0.08
	(7-9 years)	% / std. dev.	45%	2.9	10.1	0.44	12,452	0.4	3.2	0.50	0.41	0.41	0.28
	Upper middle school	mean	3,104	21.9	29.8	0.31	13,495	11.9	4.2	0.38	0.30	0.15	0.17
	(10-12 years)	% / std. dev.	26%	3.0	10.3	0.46	14,303	0.4	4.2	0.49	0.46	0.36	0.37
	College	mean	807	22.1	29.0	0.35	20,729	15.0	7.3	0.19	0.38	0.12	0.31
	(13+ years)	% / std. dev.	7%	3.2	9.2	0.48	16,000	0.9	4.7	0.39	0.49	0.32	0.46
	Whole sample	mean	12,065	21.9	31.2	0.30	11,169	9.1	2.9	0.46	0.24	0.19	0.11
		std. dev.		3.0	11.7	0.46	13,518	3.3	3.8	0.50	0.43	0.39	0.32

Note: Income is Equivalent household income (2009 ¥); Educ is education; resp. is respondent('s)

(a) Average levels of BMI

To evaluate differences in average BMI levels by childhood SES we present the results of various OLS models in Table 3.3. The first column shows our basic model, which only includes a set of control variables as regressors. There are no differences in average BMI levels between men and women, but married individuals have higher BMI. BMI also increase over time and there are significant differences between coastal and inland provinces, but not between urban and rural areas.

We initially find no evidence for an effect of maternal education, but only do so when we account for the potential rotation of the childhood SES-BMI gradient with age in the third model. Each year of additional maternal education decreases BMI with 0.128 kg/m^2, and for each year that respondents age it falls with 0.004 kg/m^2. This confirms what was shown in Figure 3.2: children of mothers with high education initially have lower BMI levels, but as they age their weight increases faster and they become on average heavier than their respondents from mothers with no or low education.

We now established that pathway 8 (childhood SES ⇨ BMI) from Figure 3.1 operates through age, but still need to explore how it relates to other pathways. In model 3 we first examine the role of pathway 10 (adulthood SES ⇨ BMI) by including the years of education of the respondent. This has no effect on BMI and does not affect the effects of maternal education. However, after introducing an interaction with age in model 5, own education explains BMI levels and maternal education is no longer significant. The correlation between maternal and own education seems to explain the relation of the former with BMI. The rotation made this unnoticeable in Table 3.2. Hence, pathway 10 in Figure 3.1 (adulthood SES ⇨ BMI) seems to be important and to explain the direct

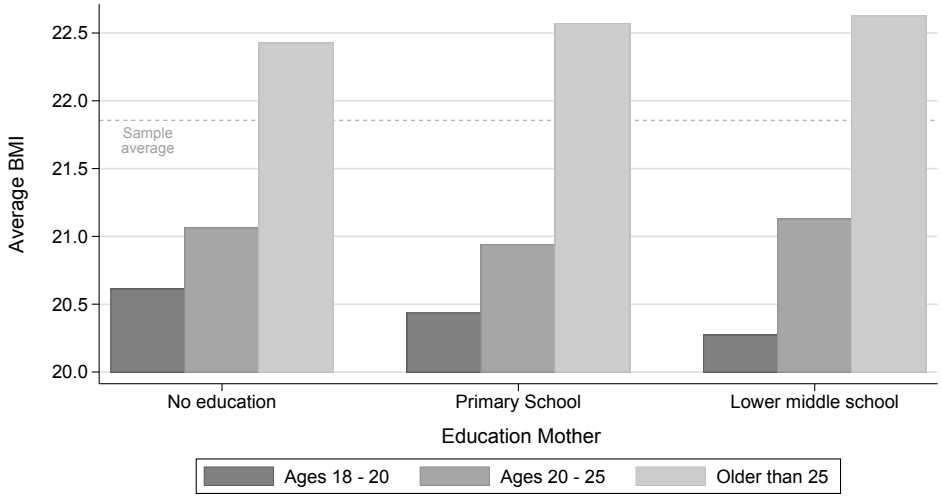


FIGURE 3.2: Average BMI by maternal education and age

TABLE 3.3: OLS regression results

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Year: 1991	0.221**	0.221**	0.220**	0.219**	0.219**	0.252**
1993	0.361***	0.363***	0.369***	0.365***	0.367***	0.429***
1997	0.598***	0.605***	0.618***	0.606***	0.623***	0.693***
2000	0.838***	0.850***	0.875***	0.857***	0.869***	0.962***
2004	0.997***	1.021***	1.051***	1.030***	1.027***	1.128***
2006	1.118***	1.144***	1.173***	1.147***	1.148***	1.261***
2009	1.241***	1.271***	1.284***	1.263***	1.260***	1.346***
Region: Coastal Urban	0.716***	0.737***	0.730***	0.695***	0.728***	0.664***
Coastal Rural	0.799***	0.800***	0.798***	0.796***	0.799***	0.786***
Inland Urban	0.200	0.210	0.215	0.195	0.210	0.191
Married Male	-0.094	-0.096	-0.087	-0.108	-0.134	0.619*
Unmarried Female	-0.741***	-0.736***	-0.725***	-0.736***	-0.753***	-0.744***
Unmarried Male	-0.565**	-0.559**	-0.544**	-0.554**	-0.568**	0.161
Age	0.170***	0.168***	0.140***	0.137***	0.091**	0.091**
Age-squared	-0.002***	-0.002***	-0.001***	-0.001**	-0.001**	-0.001**
Years of Education Mother		-0.009	-0.128**	-0.132**	-0.092	-0.091
Yrs Edu Mom X Age			0.004*	0.004*	0.003	0.003
Years of Education				0.020	-0.111**	-0.107**
Yrs Edu X Age					0.004**	0.004**
Caloric consumption females/1,000						0.406***
Caloric consumption males/1,000						0.085*
Low level of daily activity						0.358***
Constant	17.772***	17.827***	18.376***	18.272***	19.556***	18.479***
Adjusted R-squared	0.120	0.120	0.122	0.122	0.125	0.130

* $p<0.05$; ** $p<0.01$; *** $p<0.001$

pathway 8 from childhood SES to adulthood BMI. This does not explain the relationship between childhood SES and adulthood SES as this can still run through childhood health (pathway 1 and 4 (childhood SES ⇨ childhood health ⇨ adulthood SES)) or be a direct link (pathway 3 (childhood SES ⇨ adulthood SES)).

Subsequently, we assess the impact of lifestyles on BMI (pathway 7) in model 6. These lifestyles determine the level of energy consumed and expensed by an individual. We find that additional caloric consumption raises BMI for both women and men, but more so for the former than for the latter. One thousand additional calories a day will lead to a 0.406 kg/m^2 for women, but only 0.085 kg/m^2 for men. Additionally, higher levels of daily activity lead to lower BMI levels. This shows that lifestyle operate as one would expect, but it seems that it operates independently from the pathways that run through

socio-economic status either in childhood or later in life as those coefficients remain unchanged.

The final pathway (9) hypothesized in Figure 3.1 through which childhood SES could affect adulthood BMI runs through childhood health, which was proxied by early adulthood health. We explore this pathway separately because of the aforementioned data limitations. We repeat our analysis on this smaller sample and enter dummies for SRH to the regression (model 7) after we entered the variables for the respondents' own education (models 2-5). These results are included in Table A. 3.1. The effects of many of the background variables in the first model are now insignificant, but the impact of our child- and adulthood SES indicators is similar as before. The lifestyles however, have a different impact on BMI, especially physical activity, which is no longer significant in this smaller sample. Early adulthood health itself is inversely related to BMI; those in better health have higher levels of body weight and early adulthood health does not affect the role of maternal or own education. So, early adulthood health seems to have an independent effect on BMI, which is confirmed by undisclosed models that exclude these SES variables.

(b) Prevalence of overweight and obesity

The above analysis clarifies which dynamics between childhood and adulthood are important in explaining why BMI levels in China are rising. However, it does not show the resulting burden of disease as still a large share of the variation occurs within the healthy bounds of BMI in China. Consequently, not all increases in BMI are considered unhealthy. Therefore, we also estimate the effects of child- and adulthood SES on threshold values of BMI, which we define as 23 kg/m^2 and 27.5 kg/m^2, for overweight and obesity respectively.

In Table 3.4 we present the results of these linear probability models. For reasons of brevity and the similar effects of other determinants as in the earlier models, we limit ourselves here to the effects of maternal education and own education. The results of these variables are similar in the models for overweight to those on BMI levels. They also show an effect of maternal education after accounting for the rotation of the gradient with age (model 3), which dissipates once a rotating adulthood SES-BMI gradient is accounted for (model 5). For the obesity models however, the story is different. Neither maternal nor individual education has an impact on the probability of being obese even after accounting for the rotation of the gradient (models 3 and 5). Controlling for the effect of early adulthood health suggest the same conclusions[35]. The probability of being overweight is

35. We did not include the output of these models, because of the earlier described limitations of the data and the limited additional insight resulting from the models including childhood health

TABLE 3.4: Linear probabilities of being overweight (BMI > 23) or obese (BMI > 27.5)

		Model 2	Model 3	Model 4	Model 5
Overweight	Years of Education Mother	0.0003	-0.0158**	-0.0167**	-0.0098
	Yrs Edu Mom X Age		0.0006*	0.0006*	0.0004
	Years of Education			0.0045*	-0.0178***
	Yrs Edu X Age				0.0006***
Obesity	Years of Education Mother	0.0002	-0.0022	-0.0022	-0.0017
	Yrs Edu Mom X Age		0.0001	0.0001	0.0001
	Years of Education			0.0000	-0.0015
	Yrs Edu X Age				0.0000

* $p<0.05$; ** $p<0.01$; *** $p<0.001$

lower for those who experienced poor health before the age of 25, but the probability of being obese is not.

(c) The impact of child- and adulthood SES along the distribution of BMI

If the SES-BMI gradient is absent at the upper end of the BMI distribution – where BMI has most negative health outcomes – it is difficult to conclude anything about its effect on public health. Therefore, we assess the impact of child- and adulthood SES over the entire distribution of BMI using quantile regressions. The exact effects at each quantile are less important here, since we are not interested in the impact of our SES indicators at a specific BMI quantile, but more in how they compare across the distribution. We present the results graphically in Figure 3.3[36] to facilitate comparison[37].

We limit our focus to the complete model that includes both the relationship with own and maternal education and lifestyles. We again only present the effects of the SES variables on the distribution of BMI. For reference, vertical lines have been added to the plots to indicate which share of our sample is underweight, overweight and obese, i.e. have a BMI smaller than 18.5, larger than 23 or larger than 27.5 respectively. A 95% confidence interval has been added for statistical inference. The figure shows that the impact of the SES variables, both during child- and adulthood, is insignificant, and not only at the upper bound but also at the lower levels of SES[38]. However, across most of the BMI distribution all SES variables are associated with BMI and also the education effects rotate with age. The importance of maternal education is remarkable here since in none of the earlier models we found any effect of either the linear or the rotating effect after

36. Because the sequential quantile regressions are computational demanding, only the following quantiles have been plotted: 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 96, 97, 98 and 99.

37. In Table A. 3.2 the coefficients are added for all each of the estimated quantiles.

38. We checked whether this was a result of very small variations in the SES variables at the tails of the BMI distribution, but these were comparable across the entire distribution

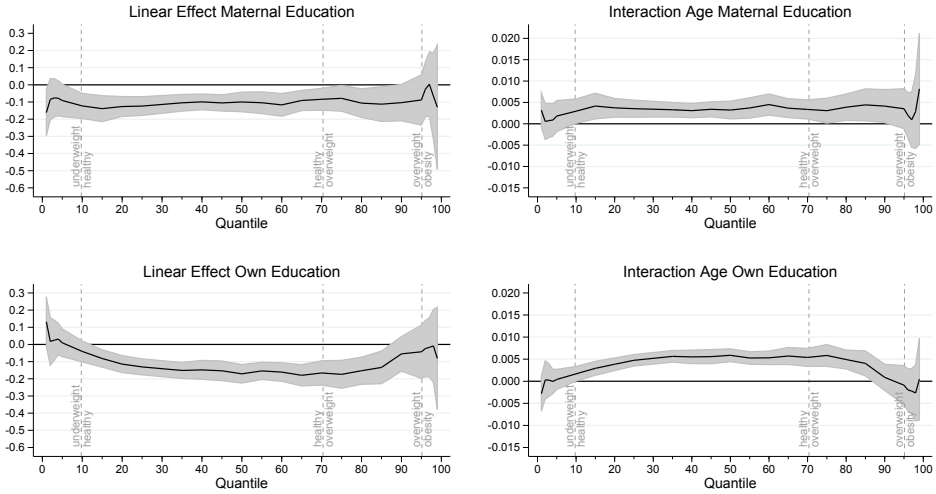


FIGURE 3.3: Effects of SES variables at different quantiles of the BMI distribution

controlling for respondents' own education. This means that pathway 8 (direct childhood SES ⇨ BMI) from Figure 3.1 could still be important in describing the shift in the BMI distribution in China. Changes in the levels of maternal education, which will most likely occur for future generations given the results in Table 3.2, might affect the BMI levels at the middle quantiles of BMI. These are currently within the healthy bounds of BMI, but may soon no longer be if the current trend continues. Consequently, socio-economic health differences might arise from this relationship not only based on current SES status, but also based on early life conditions.

3.6. DISCUSSION

The rising BMI levels in China are a concern from a public health perspective and may contribute to (future) socio-economic inequalities in health. The emerging SES gradient calls for a better understanding of what factors are driving the rise in BMI and how those factors relate to SES. We have used early life socio-economic status to preclude effects of BMI on socio-economic status. We find an association with childhood SES and discuss two possible mediating factors – childhood health and SES later in life – that might help explain it. We have shown that each of these two add to the concerns about an equitable BMI distribution in their own right. We also reveal some of the complexity behind the trend of increasing BMI levels by considering the entire distribution in addition to average levels and specific levels. We find that the SES-BMI gradient is unambiguous in the middle of the distribution, but that the proposed pathways are not as straightforward at the tails, where BMI has the largest health effects.

Our most striking result is not that SES is associated with BMI, but that this association changes with age. This is in line with the findings of Baum and Ruhm (2009) for BMI and Currie (2009) for childhood health, both in the United States. In the case of China it is in line with Ma (2010). In contrast to the results for the U.S., where the gradient widens with age, it reverses in China. At young ages low SES is associated with relatively high BMI levels, but at older ages with lower BMI. At present, especially at young ages, BMI levels for most individuals in both groups are below the threshold for overweight. While the rotating effect of maternal education disappears after controlling for own education, own education itself rotates with age as well. We cannot explain this finding. While Currie (2009) also describes a steepening of the gradient, it has not yet been observed that this trend continues this far into adulthood. One possible explanation is that higher educated mothers are more concerned about their children's weight, but that higher educated individuals are less concerned about their own weight and seize full opportunity to consume food and go without physical activity once those choices are up to themselves. The CHNS data unfortunately does not provide sufficiently detailed information to investigate this explanation any further.

There are other possibilities that could explain the rotation. First, it is possible that the rotation with age is in fact capturing something that is not included in our model, such as changes in education itself or effects of education that diminish over time. For example, if the quality of education improves over time, education may be less beneficial for older respondents. Second, the main age effect may have been misspecified. If the linear and squared age effect do not capture all age effects and age is related to maternal and/or own education levels, the rotation might be measuring these uncaptured effects. We therefore investigated the role of age further, through higher age polynomials and age dummies, but these did not explain the rotation of education effects with age. A third possible explanation is that there is a differential effect of education for men and women, which is captured in the rotation because in our sample adult men are on average 4.5 years older than women. Consequently, a difference in education effects on BMI for men and women could have been captured in the interaction between age and education. We also looked into this explanation and found a significant difference in education effects for men and women, which is remarkable given the insignificance of the gender dummy in the models. Inclusion of the interaction, however, only marginally changes the current model estimates. In stratified models by gender, we also still find a rotation of the SES-BMI gradient with age, albeit less strong for males. This could be a result of the small sample sizes in the stratified models. A final explanation might be that the SES-BMI gradient differs between regions, but this also was not supported by stratified models.

A second striking result is the fact that a SES-BMI gradient is only evident in the middle of the BMI distribution and not at its tails. This raises the question whether we should be concerned about the fairness of the BMI distribution. Although this also holds for BMI

levels that are below the obesity threshold, it does not for the prevalence of overweight. Overweight is an important determinant of future health outcomes, albeit not as serious as obesity. What might be more worrying is the overall trend in average BMI levels in China. If the average level of BMI keeps rising at the average pace of our study period, it may not take much longer before the quantiles at which SES affects BMI no longer represent healthy levels of BMI. While we can only speculate about future BMI development in China, current developments in relation to SES suggest that the worst is yet to come.

Our findings bear some resemblance to those of Ma (2010) who also investigated the SES-BMI gradient using maternal education in China. However, we examined health outcomes later in life and addressed potential pathways, especially those through health levels at younger ages. Our results can be compared to Shankar (2009) who ran quantile regression models on the 2006 wave of the CHNS data, but found little evidence for any SES gradient across the entire distribution. The discrepancy with our findings is likely to be the result of the inclusion of only a single wave of the CHNS, to not restricting the analysis to those who had lived with their mother throughout the study period and to not accounting for the rotation of the gradient.

The CHNS data used throughout have their strengths and weaknesses. On the one hand, it is unique in stretching across a very large and important time period and including detailed information on both diets and body measurements, on the other hand it is also limited in some other respects. The linkage required between mothers and children limited the number of usable observations. Especially at older ages, the number of mothers that could be linked to their children was small and may not be unrelated to SES, since mothers with lower SES may be more likely to live with their children.

In conclusion, China is facing a substantial rise in BMI levels, which is unlikely to be identical by socio-economic status. The expected impact of the SES differential is not straightforward, as it is yet unclear whether it will be the higher or the lower SES groups that face of the more negative health consequences from high body mass. This study cannot yet answer that question. This is because it cannot claim causality of the findings. It has addressed some endogeneity concerns by focusing on maternal education, but more is probably necessary for a full comprehension of these dynamics. Further research should focus on why SES is related to BMI and what causes the gradient to rotate. Only a deeper understanding of these mechanisms will provide the insights required to successfully curtail the current developments in BMI and the resulting negative health outcomes.

TABLE A. 3.1: Pooled OLS for subsample including early adulthood health

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Year: 1991	-0.304	-0.276	-0.301	-0.317	-0.349	-0.514*	-0.493*
1993	-0.303	-0.277	-0.287	-0.302	-0.325	-0.453	-0.430
1997	-0.337	-0.315	-0.306	-0.314	-0.315	-0.372	-0.348
2000	-0.136	-0.120	-0.100	-0.104	-0.106	-0.133	-0.121
2004	-0.142	-0.136	-0.105	-0.109	-0.110	-0.110	-0.109
2006	-0.024	-0.020	0.007	0.009	0.002	0.034	0.042
Region: Coastal Urban	0.569**	0.559**	0.561**	0.588**	0.600**	0.627**	0.633**
Coastal Rural	0.776***	0.777***	0.782***	0.784***	0.793***	0.787***	0.728***
Inland Urban	0.100	0.094	0.099	0.114	0.134	0.162	0.174
Married Male	0.079	0.080	0.083	0.087	0.067	0.906**	0.886*
Unmarried Female	-0.681**	-0.686**	-0.707**	-0.699**	-0.703**	-0.665**	-0.672**
Unmarried Male	-0.366	-0.371	-0.381	-0.375	-0.371	0.478	0.465
Age	0.159*	0.159*	0.092	0.097	0.026	0.028	0.033
Age-squared	-0.001	-0.001	0.000	0.000	0.000	0.000	0.000
Years of Education Mother		0.005	-0.123	-0.121	-0.067	-0.067	-0.062
Yrs Edu Mom X Age			0.006	0.006	0.003	0.003	0.003
Years of Education				-0.013	-0.218**	-0.188*	-0.194*
Yrs Edu X Age					0.008*	0.008*	0.008*
Caloric consumption females/1,000						0.576***	0.557***
Caloric consumption males/1,000						0.163*	0.142*
Low level of daily activity						0.019	0.020
Early adulthood SAH was Good							-0.412**
Early adulthood SAH was Fair							-0.565**
Early adulthood SAH was Poor							-1.591**
Constant		18.287***	19.351***	19.398***	21.111***	19.665***	20.054***
Adjusted R-squared		0.109	0.110	0.110	0.111	0.119	0.124

TABLE A. 3.2: Quantile regression estimates for selected quantiles

	q01	q02	q03	q04	q05	q10
1991	-0.077	0.040	-0.044	-0.011	0.042	0.086
1993	0.230	0.229	0.155	0.220	0.208	0.220
1997	0.133	0.182	0.163	0.170	0.154	0.316*
2000	0.081	0.235	0.128	0.205	0.278	0.408**
2004	0.079	0.282	0.124	0.044	0.122	0.347*
2006	0.064	-0.015	-0.008	0.157	0.277	0.478**
2009	-0.123	0.015	-0.044	0.089	0.209	0.368*
Region: Coastal Urban	0.263	0.226	0.208	0.140	0.169	0.137
Coastal Rural	0.355	0.116	0.218	0.270*	0.245*	0.259*
Inland Urban	-0.007	-0.125	-0.090	-0.104	-0.132	-0.002
Married Male	-0.311	-0.320	-0.071	0.004	0.066	0.172
Unmarried Female	-0.413	-0.430*	-0.426*	-0.429*	-0.444**	-0.485***
Unmarried Male	-0.390	-0.471	-0.139	-0.081	-0.067	-0.136
Age	0.176***	0.134**	0.129**	0.124***	0.101**	0.075**
Age-squared	-0.002***	-0.002**	-0.002**	-0.001***	-0.001***	-0.001***
Years of Education Mother	-0.161*	-0.084	-0.076	-0.078	-0.091	-0.123**
Yrs Edu Mom X Age	0.003	0.001	0.001	0.001	0.002	0.003*
Years of Education	0.132	0.018	0.025	0.031	0.010	-0.041
Yrs Edu X Age	-0.003	0.000	0.000	-0.000	0.000	0.002*
Caloric consumption females	0.021	0.026	0.163	0.202*	0.242**	0.257***
Caloric consumption males	-0.028	-0.009	0.022	0.035	0.049	0.072
Low level of daily activity	-0.166	-0.234*	-0.220*	-0.218*	-0.191*	-0.054
Constant	13.531***	15.200***	15.072***	15.193***	15.698***	16.903***

TABLE A. 3.2: Quantile regression estimates for selected quantiles (cont.)

	q15	q20	q25	q30	q35	q40
1991	0.102	0.137	0.150	0.080	0.101	0.104
1993	0.161	0.238*	0.262*	0.233	0.275*	0.248**
1997	0.345*	0.310**	0.338**	0.240	0.287*	0.308**
2000	0.488***	0.475***	0.529***	0.449	0.531***	0.614***
2004	0.451***	0.517***	0.579***	0.613	0.651***	0.712***
2006	0.442**	0.550***	0.716***	0.691	0.695***	0.811***
2009	0.448**	0.550***	0.693***	0.662	0.767***	0.829***
Region: Coastal Urban	0.221	0.329**	0.347***	0.433	0.517***	0.537***
Coastal Rural	0.380***	0.418***	0.487***	0.595	0.662***	0.734***
Inland Urban	-0.034	-0.044	0.013	0.037	0.107	0.063
Married Male	0.310	0.405	0.470*	0.726	0.821***	0.914***
Unmarried Female	-0.473***	-0.391**	-0.429***	-0.320	-0.326**	-0.296*
Unmarried Male	-0.080	0.073	0.173	0.435	0.438	0.543*
Age	0.059*	0.065**	0.060**	0.077	0.062**	0.070***
Age-squared	-0.001**	-0.001***	-0.001***	-0.001	-0.001***	-0.001***
Years of Education Mother	-0.139***	-0.126***	-0.124***	-0.112	-0.105***	-0.099***
Yrs Edu Mom X Age	0.004**	0.004***	0.004***	0.003	0.003***	0.003***
Years of Education	-0.082**	-0.114***	-0.131***	-0.128	-0.151***	-0.148***
Yrs Edu X Age	0.003***	0.004***	0.005***	0.005	0.006***	0.006***
Caloric consumption females	0.258***	0.324***	0.329***	0.370	0.389***	0.410***
Caloric consumption males	0.059	0.100	0.103	0.094	0.115*	0.092*
Low level of daily activity	0.015	0.116	0.104	0.105	0.140*	0.156*
Constant	17.622***	17.663***	17.894***	17.611	18.031***	18.044***

TABLE A. 3.2: Quantile regression estimates for selected quantiles (cont.)

	q45	q50	q55	q60	q65	q70
1991	0.149	0.182	0.213*	0.230*	0.257*	0.315*
1993	0.221*	0.259*	0.343**	0.468***	0.493***	0.487***
1997	0.371**	0.466***	0.585***	0.673***	0.690***	0.772***
2000	0.667***	0.756***	0.843***	0.993***	1.102***	1.170***
2004	0.828***	0.925***	1.074***	1.242***	1.380***	1.461***
2006	0.865***	0.970***	1.110***	1.280***	1.512***	1.566***
2009	0.891***	1.039***	1.169***	1.372***	1.570***	1.743***
Region: Coastal Urban	0.615***	0.652***	0.657***	0.711***	0.704***	0.788***
Coastal Rural	0.779***	0.779***	0.810***	0.864***	0.947***	0.970***
Inland Urban	0.114	0.125	0.138	0.165*	0.192*	0.205*
Married Male	1.024***	1.037***	1.277***	1.222***	1.165***	0.933**
Unmarried Female	-0.333*	-0.469***	-0.429**	-0.642***	-0.805***	-0.880***
Unmarried Male	0.619*	0.607*	0.816**	0.778**	0.700*	0.444
Age	0.075***	0.061**	0.084***	0.081***	0.064*	0.071**
Age-squared	-0.001***	-0.001***	-0.001***	-0.001***	-0.001**	-0.001**
Years of Education Mother	-0.105***	-0.100***	-0.104**	-0.117***	-0.091**	-0.084**
Yrs Edu Mom X Age	0.003***	0.003**	0.004**	0.005***	0.004**	0.003**
Years of Education	-0.153***	-0.171***	-0.154***	-0.161***	-0.180***	-0.166***
Yrs Edu X Age	0.006***	0.006***	0.005***	0.005***	0.006***	0.005***
Caloric consumption females	0.486***	0.506***	0.590***	0.675***	0.708***	0.645***
Caloric consumption males	0.096*	0.088*	0.070	0.085	0.077	0.083
Low level of daily activity	0.202**	0.234**	0.277**	0.343***	0.392***	0.452***
Constant	18.076***	18.639***	18.213***	18.467***	19.101***	19.358***

TABLE A. 3.2: Quantile regression estimates for selected quantiles (cont.)

	q75	q80	q85	q90	q95	q96
1991	0.383**	0.374*	0.297	0.592**	0.631**	0.771**
1993	0.614***	0.609***	0.620**	0.649**	0.784**	0.825**
1997	0.870***	0.921***	1.065***	1.335***	1.658***	1.807***
2000	1.306***	1.416***	1.464***	1.820***	2.279***	2.621***
2004	1.567***	1.682***	1.757***	1.995***	2.421***	2.579***
2006	1.810***	2.026***	2.144***	2.297***	2.913***	3.200***
2009	1.937***	2.136***	2.357***	2.717***	3.093***	3.210***
Region: Coastal Urban	0.836***	0.963***	1.038***	1.244***	1.177***	1.220***
Coastal Rural	0.966***	1.143***	1.206***	1.159***	1.248***	1.321***
Inland Urban	0.233*	0.330**	0.466***	0.563***	0.740***	0.859***
Married Male	0.962***	1.134***	1.127**	1.066*	0.403	0.558
Unmarried Female	-0.833***	-0.872***	-0.938***	-1.027***	-1.639***	-1.626***
Unmarried Male	0.468	0.536	0.392	0.211	-0.636	-0.337
Age	0.062*	0.069*	0.092**	0.127***	0.141*	0.156*
Age-squared	-0.001*	-0.001*	-0.001**	-0.001*	-0.001	-0.001
Years of Education Mother	-0.078*	-0.106*	-0.112*	-0.104	-0.088	-0.025
Yrs Edu Mom X Age	0.003*	0.004*	0.004*	0.004*	0.004	0.002
Years of Education	-0.174***	-0.153***	-0.134**	-0.055	-0.043	-0.024
Yrs Edu X Age	0.006***	0.005***	0.004**	0.001	-0.001	-0.002
Caloric consumption females	0.632***	0.651***	0.737***	0.717***	0.504**	0.574**
Caloric consumption males	0.069	0.058	0.148	0.169*	0.162	0.122
Low level of daily activity	0.514***	0.581***	0.753***	0.875***	0.907***	0.914***
Constant	19.717***	19.887***	19.769***	19.593***	21.390***	21.103***

TABLE A. 3.2: Quantile regression estimates for selected quantiles (cont.)

	q97	q98	q99
1991	0.852**	0.695	0.813
1993	0.899**	1.022**	1.210*
1997	1.928***	2.211***	2.501***
2000	2.622***	2.968***	3.033***
2004	2.728***	3.124***	3.707***
2006	3.393***	4.019***	4.454***
2009	3.392***	4.140***	3.909***
Region: Coastal Urban	1.053***	1.005**	1.048*
Coastal Rural	1.291***	1.529***	1.735***
Inland Urban	0.940***	1.168**	1.764***
Married Male	0.523	1.007	0.706
Unmarried Female	-1.492***	-1.822**	-2.746**
Unmarried Male	-0.402	0.373	-0.085
Age	0.138	0.080	-0.150
Age-squared	-0.001	-0.000	0.003
Years of Education Mother	0.004	-0.065	-0.130
Yrs Edu Mom X Age	0.001	0.003	0.008
Years of Education	-0.017	-0.008	-0.081
Yrs Edu X Age	-0.002	-0.003	0.000
Caloric consumption females	0.441	0.646*	0.708
Caloric consumption males	0.136	0.065	-0.010
Low level of daily activity	0.824***	0.859**	0.387
Constant	21.744***	22.705***	27.820***

* $p<0.05$; ** $p<0.01$; *** $p<0.001$

Chapter 4

Rising inequalities in income and health in China: Who is left behind?

In collaboration with: Tom Van Ourti and Eddy van Doorslaer

ABSTRACT

Background:

During the last decades, China has experienced double-digit economic growth rates and rising inequality.

Methods:

This paper implements a new decomposition on the China Health and Nutrition panel Survey (1991-2006) to examine the extent to which changes in level and distribution of incomes and in income mobility are related to health disparities between rich and poor.

Results:

We find that health disparities in China relate to rising income inequality and in particular to the adverse health and income experience of older (wo)men, but not to the growth rate of average incomes over the last decades.

Conclusion:

These findings suggest that replacement incomes and pensions at older ages may be one of the most important policy levers in combating health disparities between rich and poor Chinese.

4.1. INTRODUCTION

The relationship between income and health, and the inequality in each of these, has been documented in many settings, but few countries have experienced changes in their income distribution as dramatic and as rapid as China has over the last few decades. China's rapid transition since the early 1980s from a completely planned economy to a more market led economy has led to unprecedented economic growth: China's real GDP in 2009 was more than the twelvefold of its real GDP in 1980. Poverty reductions were similarly spectacular: between 1981 and 2005 the poverty headcount ratio ($1.25 poverty line) fell from 94% to 26% in rural areas and from 44% to 2% in urban areas (World Bank, 2005). The income distribution has not only shifted upward, it has also substantially widened, leading to rising inequality in incomes (Benjamin et al., 2008) both within and between regions.

However, this rapid economic growth did not spark equally impressive health improvements (Tang et al., 2008). While, for instance, average life expectancy has continued to grow, China has lost its high achiever position. The Chinese average life expectancy is not any longer higher than expected on the basis of average income, indicating that the pace of health gains has slowed down. Moreover, health disparities between urban and rural regions and between rich and poor have grown and have given rise to increased public dissatisfaction. The Chinese government has recognized these challenges and has responded to them. Reform of health care has been ongoing for several years and in 2008 the Ministry of Health announced major new policy directions for achieving Healthy China by 2020 (Tang et al., 2008). These reforms have somewhat alleviated the health burdens faced by the (rural) poor, but concerns remain (Hougaard et al., 2011; Wagstaff et al., 2009b). At the same time, China has experienced several other demographic and economic transitions. The population is rapidly growing older, becoming more urbanized and economic activity has shifted from largely agricultural to industrial. A priori, it is not obvious how these dramatic changes in the Chinese society have influenced the relation between (changes in) income, health, their distributions and the health gaps between rich and poor.

One way of uncovering the mechanics underlying these relationships is to decompose the (changes in the) degree of income-related health inequality (IRHI). Contoyannis and Forster (1999) initiated the examination of conditions under which increases in income and income inequality will be associated with higher levels of IRHI as measured by a rank-based concentration index. Van Ourti et al. (2009) have generalized their argument using a cohort decomposition approach and have used it to examine inequalities in health by income in European countries in the 1990s. In this paper, we propose an extension and simplification of their decomposition framework to uncover these mechanics in China. The simplification derives from recognizing that health – unlike income – is a bounded variable at the upper end and using an adjusted rank-based measure of absolute inequal-

ity. The extension consists of the addition of an additional component relating to income mobility which proves to be crucial for an improved understanding of the evolution of IRHI. We then apply this framework to an analysis of inequalities in self-reported health by income in a cohort followed over 6 waves of the China Health and Nutrition Survey (CHNS) panel and spanning a period of 15 years (1991-2006).

We find that the degree to which health is associated to income – as measured by a rank-based measure of absolute inequality – more than tripled over those 15 years. This substantive rise turns out to be not related to the massive income growth and only weakly to the growth in income inequality. It is almost entirely explained by the very different income mobility by age and gender. Especially the much more adverse experience of older females in terms of their income position and health seems to have been important. The absence of adequate replacement incomes for sick and old women, especially in rural areas, is a likely candidate to explain this phenomenon.

In the next section we briefly review what is known on income and health developments in China since the start of the economic reforms. In section 4.3 we discuss our new decomposition framework which is then applied to Chinese panel data spanning the period 1991-2006 in section 4.4. We conclude with a discussion of potential implications.

4.2. INCOME, HEALTH AND INEQUALITY IN CHINA

China is no different from any other country in exhibiting inequalities in income and health and a strong association between the two. What makes China more unique is (i) the pace of changes in recent decades and (ii) the sheer size of the country reflecting such immense differences in levels of development. The rapid growth in average incomes in China has been accompanied by an increase in income inequality. While initially, income inequality was at a fairly low level (Kanbur and Zhang, 2005), the transition from a planned to a more market oriented economy has dramatically increased variability in incomes (Gustafsson et al., 2008; Lin et al., 2008; Shen and Yao, 2008; Sutherland and Yao, 2011; Yang, 1999), wealth (Gustafsson et al., 2006) and consumption (Keidel, 2009). Initially, rising income inequality was not viewed as unacceptable, because the initial income distribution was relatively egalitarian and because it reduced absolute levels of poverty in society. Moreover, high income mobility was perceived as a blessing because it meant that poor households were not trapped in their situation (Khor and Pencavel, 2010), but more recently discontent about the widening gaps has been growing (Yu, 2008).

Developments in nationwide income inequality obviously mask important geographic differentials. Two geographic distinctions appear particularly relevant in this respect. A first important geographic division of China is that between eastern provinces near the coast and the western non-coast provinces. Historically the eastern coastal regions have had higher levels of GDP than the interior regions (Chen and Fleisher, 1996). More

recently, the coastal provinces have coupled a stronger rise in mean per capita income with a more modest increase in income inequality compared to the interior regions (Benjamin et al., 2008). While the growth of the inter-regional inequality was anticipated by the government, which favored the coastal provinces in the early years of the reforms, the spillovers to other regions were insufficient to reduce the initial inter-regional income inequality (Brun et al., 2002).

A second geographic divide is that between rural and urban areas. Large urban-rural disparities in income have persisted for a long time, among others due to the introduction of the labor-mobility restricting Hukou system (Wang and Piesse, 2010), and to welfare and financial policies that favored the urban areas (Yang, 1999). This inequality between rural and urban areas contributes substantially to total income inequality in China, and is rapidly growing (Yang, 1999) such that China's urban-rural income gap is now one of the largest in the world (Chang, 2002; Sicular et al., 2007). Within-urban and within-rural income inequality are currently of similar magnitude (Sutherland and Yao, 2011).

Several studies have documented health variation by income and region in China. For example, (Liu et al., 2008) report an income gradient in self-reported health in rural areas, while (Chen and Meltzer, 2008) report income gradients for hypertension and obesity. Other studies have documented substantial health variation by region, e.g. life expectancy between and within provinces (Tang et al., 2008), infant mortality between and within inland and coastal regions (Zhang and Kanbur, 2005) and between rural and urban areas (Shi et al., 2008), or self-reported health between rural and urban areas (Van de Poel et al., 2012).

Fewer studies have measured trends in IRHI. Chen et al. (2007) use the China Health and Nutrition Survey to find that child malnutrition was concentrated among the poor in 1989, that this concentration decreased in the early nineties but rose again in 2000. Using the same data, Yip (2010) does not find these inequalities to have either widened or narrowed. Wagstaff et al. (2009a) use data from the maternal and child health (MCH) surveillance system of 2003 and report large inequalities in infant, under-five and maternal deaths across counties. Recently, Feng et al. (2010) examined the evolution of socio-economic disparities in maternal mortality by county (ranked by income) between 1996 and 2006 in the MCH. They also do not find a clear trend in the concentration indices.

Overall, we conclude that (i) there is enormous variation in China in both the level and distribution of income and health, (ii) which are associated across place and time and (iii) no study has yet attempted to trace out how developments of income growth, income inequality and income mobility are related to income-related health inequality developments over the recent decades.

4.3. DECOMPOSING INCOME-RELATED HEALTH INEQUALITIES INTO INCOME GROWTH, MEAN-PRESERVING INCOME CHANGES AND INCOME MOBILITY

We measure income-related health inequalities by the variation of health across the income dimension (Erreygers, 2009a; Wagstaff and Doorslaer, 2000) and summarize its association with the income distribution using a decomposition methodology (Wagstaff et al., 2003). Since this paper explicitly focuses on the exceptional income growth rates that China experienced over the last three decades, we isolate the role of changes in average incomes from other changes in the income distribution. To this end, we extend the decomposition method proposed by Van Ourti et al. (2009) such that it accounts for the bounded nature of health (Erreygers and Van Ourti, 2011b). We also provide a new interpretation of this decomposition that stresses the role of income mobility.

(a) Measurement of income-related health inequalities

IRHI is most commonly measured using the standard concentration index. It resembles a Gini index of health, but replaces the rank of health by the rank of income such that it measures the variation of health across the income dimension (Wagstaff et al., 1991). An important implicit value judgment is that equi-proportionate health changes leave the concentration index unchanged. This is a common assumption in the uni-dimensional income inequality literature (Lambert, 2001), but recent work shows that this focus on relative health differences is less innocuous in the case of bi-dimensional IRHI (Erreygers, 2009a, 2009b; Erreygers and Van Ourti, 2011a, 2011b; Lambert and Zheng, 2011; Wagstaff, 2009, 2011b, 2011a). When health has a finite upper and lower bound (as in our empirical application) an impossibility result arises: relative health differences are incompatible with the requirement that income-related inequalities in *health* and *ill-health* should rank a set of health-income distributions in the same way. We put more emphasis on the latter requirement since the transformation from health to ill-health, i.e. ill-health equals maximum health minus actual health, is arbitrary and does not change the information contained in the health (or ill-health) variable. It follows that we can no longer resort to the concentration index that focuses on relative health differences, but instead use the Erreygers index (2009a) which indicates that IRHI remains unchanged under equal health additions:

$$\text{eq 4.1} \qquad C_E\left(h\,|\,y\right)=\frac{8}{n^2\left(h^{\max}-h^{\min}\right)}\sum_{i=1}^{n} z_i h_i=\frac{8}{n^2}\sum_{i=1}^{n} z_i h_i$$

where h_i equals the level of health of individual i[39], y_i equals income, and n stands for the number of observations. z_i equals the deviation of individuals i's income rank from the

39. We assume in eq 4.1 without loss of generality that health is bounded between 0 ($h^{\min}$) and 1 ($h^{\max}$) as any bounded variable can be retransformed to the unit interval.

mean income rank[40]; and increases linearly between $(1-n)/2$ and $(n-1)/2$ and takes zero for the individual with the mean income rank. In other words, the z_i's are negative for the poorest half of the population and positive for the richest half, and sum to zero across the entire population. Hence, the Erreygers index will be negative if the poor have overall better health than the rich (pro-poor IRHI), and positive values will imply the opposite (pro-rich IRHI).

(b) Changes in the income distribution and IRHI: a decomposition approach

Our approach extends the approach of Van Ourti et al. (2009) in two ways. First, we simplify their decomposition by measuring IRHI with the Erreygers index. This simplification is the logical consequence of the change of focus from relative to absolute health differences, as it no longer requires to understand how changes in the income distribution are related to average health levels, which is indispensable to understand relative health differences. Second, we show that the decomposition is richer than originally envisaged by the authors. Van Ourti et al. (2009) isolate the role of changes in average incomes from other changes in the income distribution, but we show that this implies that IRHI also depend on the association between income mobility and the evolution in other determinants of health.

Our extended decomposition is a cohort-decomposition, i.e. it unravels how IRHI evolves over time for a given cohort of individuals such that we can abstract from changes in IRHI that are driven by compositional changes of the underlying population. We show later in this section that our cohort-decomposition also isolates the ageing-effect of our cohort such that we can study the change of IRHI over time in isolation from ageing. We assume in this section that individuals remain in the cohort over the entire time period, but we allow for drop-out in the empirical part of the paper. In the remainder of this section, we discuss our extended decomposition in more detail.

We start by describing the association between health and income y_i conditional on a set of K other variables x_i (e.g. age, sex) and add a time subscript $t=1,\ldots,T$.

$$\text{eq 4.2} \quad h_{it} = \alpha + \varphi(y_{it}) + x_{it}'\beta$$

$\varphi(\cdot)$ is a (non-linear) function of income, α is a parameter, and β is a parameter vector of dimension K. As the shape of the function $\varphi(\cdot)$ will largely determine the association between the evolution of IRHI and changes in the income distribution (see below), we allow for a very general functional form.

---

40. $z_i = \frac{n+1}{2} - \lambda_i$, and the income rank λ_i takes the value 1 for the richest individual and n for the poorest individual.

Three simplifying assumptions underlie eq 4.2. In the empirical part of the paper, we confirm that these assumptions are justified. First, we rule out interactions between the non-linear function of income $\varphi(\cdot)$ and the set of other variables x_i. Second, we assume that the coefficients in eq 4.2 are fixed over time which has the advantage that changes in IRHI are associated to changes in the explanatory variables only (and not to changes in the coefficients). Finally, eq 4.2 ignores dynamics. Note also that we exclude a residual term from eq 4.2 for ease of exposition.

We express the evolution of IRHI as the difference between the Erreygers index in period t and the first period, and substitute eq 4.2 in eq 4.1:

$$\text{eq 4.3} \quad C_E(h_t|\, y_t) - C_E(h_1|\, y_1) = \frac{8}{n^2}\sum_{i=1}^{n}\left\{\left[z_{it}\phi(y_{it}) - z_{i1}\phi(y_{i1})\right] + \sum_{k=1}^{K}\beta^k\left[z_{it}x_{it}^k - z_{i1}x_{i1}^k\right]\right\}$$

Eq 4.3 provides us with a decomposition of the evolution of IRHI into those related to changes in the income distribution (i.e. the first term between square brackets) and into those due to changes in other determinants, weighted by the respective z_i's (i.e. the second term between square brackets); it does not, however, allow to separate the effect of proportional income growth from other changes in the income distribution. Van Ourti et al. (2009) disentangle both effects by introducing two hypothetical health states in period t. The first is the situation of 'proportional income growth' (*pg*) in which total income growth benefits everyone proportionately such that individual incomes grow at the same rate as average income (y_t), i.e. $h_{it}^{pg} = \alpha + \phi(y_{it}^{pg}) + x_{it}'\beta$ with $y_{it}^{pg} = y_{i1}(Y_t/Y_1)$. Secondly, we introduce a health state in which income 'does not grow' (*ng*), but the other variables evolve at their actual rate, i.e. $h_{it}^{ng} = \alpha + \phi(y_{i1}) + x_{it}'\beta$.[41]

We have now sufficient information to derive our extended decomposition:

$$\text{eq 4.4} \quad C_E(h_t|\, y_t) - C_E(h_1|\, y_1) = \underbrace{\left[C_E(h_t|\, y_t) - C_E\left(h_t^{pg}|\, y_1\right)\right]}_{(2)\ \&\ (3)} + \underbrace{\left[C_E\left(h_t^{pg}|\, y_1\right) - C_E\left(h_t^{ng}|\, y_1\right)\right]}_{(1)}$$

$$+ \underbrace{\left[C_E\left(h_t^{ng}|\, y_1\right) - C_E(h_1|\, y_1)\right]}_{(4)}$$

41. Note that the introduction of the hypothetical health states is 'path-independent', i.e. it does not matter for our decomposition whether we have the sequence $h_{i1} \to h_{it}^{ng} \to h_{it}^{pg} \to h_{it}$ or $h_{i1} \to h_{i1}^{pg} \to h_{it}^{pg} \to h_{it}$, since we combine the Erreygers index – which only focuses on absolute health differences – with eq 4.2 – which is additively separable in $\varphi(y_{it})$ and $x_{it}'\beta$.

$$= \frac{8}{n^2}\left\{ \begin{array}{l} \underbrace{\sum_{i=1}^{n} z_{i1}\left[\phi\left(y_{it}^{pg}\right)-\phi\left(y_{i1}\right)\right]}_{(1)} + \underbrace{\sum_{i=1}^{n}\left[z_{it}\phi\left(y_{it}\right)-z_{i1}\phi\left(y_{it}^{pg}\right)\right]}_{(2)} \\ + \underbrace{\sum_{i=1}^{n}\left(z_{it}-z_{i1}\right)\left(\sum_{k=1}^{K}\beta^{k}x_{it}^{k}\right)}_{(3)} + \underbrace{\sum_{i=1}^{n} z_{i1}\left[\sum_{k=1}^{K}\beta^{k}\left(x_{it}^{k}-x_{i1}^{k}\right)\right]}_{(4)} \end{array} \right\}$$

which exploits that $C_E(h_t^{pg} \mid y_t^{pg}) \equiv C_E(h_t^{pg} \mid y_1)$ and $y_{it}^{ng} \equiv y_{i1}$. Eq 4.4 shows that we can decompose the evolution of IRHI into four effects: (i) a term related to income growth; (ii) a term related to other mean-preserving changes in the income distribution; (iii) a term related to differences in income mobility across non-income variables; and (iv) a term related to the association between changes in the other variables and z_i in the first period.

(i) Income growth

Term (1) in eq 4.4 captures the association between the evolution of IRHI and average income growth. It describes the difference between IRHI in the hypothetical health state in which all individuals would have had their incomes changed proportionately and IRHI in the state in which incomes would have remained at the level of the first period. Recalling that the z_{it}'s sum to zero and increase linearly when moving from the poorest to the richest individual is sufficient to understand that a positive/negative value of term (1) will occur when $\phi(y_{it}^{pg}) - \phi(y_{i1})$ is 'on average' increasing/decreasing with income. Intuitively this means that IRHI will rise/decrease when the same proportional income change has a larger/smaller health effect for individuals with a higher initial income.[42] Whether and how this relationship between health and proportional income increases varies with income will depend on the shape of $\varphi(\cdot)$. This highlights the importance of choosing a flexible functional form for income in eq 4.2, rather than – as is usual in the literature – sticking to a functional form (like e.g. the natural logarithm) that predetermines how the health effect of proportional income changes varies with income.

(ii) Other mean-preserving changes in the income distribution

Term (2) eq 4.4 measures the change in IRHI associated with changes in the distribution of income that are unrelated to proportional income growth. It combines changes from z_{i1} to z_{it} (which are a function of the income ranks) with mean-preserving changes in the income levels from $y^{pg}{}_{it}$ to y_{it}. When the income ranks do not change over time, the

42. $\phi(y_{it}^{pg}) - \phi(y_{i1})$ will only 'on average' increase/decrease with income when its partial derivative with respect to y_{i1} is positive, i.e. when

$$\partial\left[\phi\left(y_{it}^{pg}\right)-\phi\left(y_{i1}\right)\right]\Big/\partial y_{i1} \begin{array}{c}\geq\\=\\\leq\end{array} 0 \Leftrightarrow \left[\partial\phi\left(y_{it}^{pg}\right)\Big/\partial y_{it}^{pg}\right] Y_t \begin{array}{c}\geq\\=\\\leq\end{array} \left[\partial\phi\left(y_{i1}\right)\Big/\partial y_{i1}\right] Y_1 .$$

health impact of the change in income levels will depend on the functional form of $\varphi(\cdot)$ and on the extent to which income gains and losses differ across income groups. When richer individuals have higher income gains than poorer individuals (a pro-rich change) *and* when proportional income increases have a larger health effect for individuals with higher incomes, IRHI will rise (and vice versa). When both effects have the opposite sign (e.g. higher health effect of a proportional income change for richer individuals combined with a pro-poor change), it is an empirical question which effect dominates. In practice, income ranks will also change, but we cannot *a priori* predict the effect on IRHI without knowing the realized changes in income ranks. However, it seems likely that the change in income ranks will in most cases be independent or coincide with the pro-poor- or pro-richness of the change of income levels (but generally not be in the opposite direction).[43]

(iii) Income mobility across non-income variables

Term (3) in eq 4.4 illustrates that the evolution in IRHI is also driven by the association between the non-income variables in eq 4.2 and changes in the income ranks. Its interpretation depends on the sign of β^k, which measures the health effect of the non-income variables. When this effect is positive, all that is needed to increase IRHI is that those already experiencing better health because of the non-income variables, e.g. the young, move up in the income distribution, i.e. $z_{it} > z_{i1}$. When the health effect is negative, the opposite result holds. In other words, this term identifies the contribution of differences in income mobility across the non-income variables to IRHI (Allanson et al., 2010; Jones and Nicolás, 2004).

Similar to other decomposition methods (Fortin et al., 2011), the interpretation of term (3) is more complicated when x_{it} includes dummy variables representing categorical variables. First, $\beta^k x^k_{it}$ measures in that case the health effect of a change in the dummy from 0 to 1. Second, the choice of the reference category (e.g. whether we take males or females as the reference category) matters for the contribution of each dummy variable, but does

43. An implicit assumption so far has been that we conceive IRHI as being invariant to *equal* health additions and the subdivision of changes in the income distribution along a relative concept (i.e. *proportional* income growth versus 'other mean-preserving changes'). Our choice for *equal* health additions is the logical consequence of using the Erreygers index (see also before), and we find it also very plausible to define income growth as a relative concept. However, those who favor absolute income inequality concepts might find a definition of income growth where every individual gains the same *amount* of income more plausible. The sum of term (1) and (2) will not be affected by this alternative assumption and it is straightforward to adapt terms (1) and (2).

not matter for the sum across all dummy variables.[44] This means that the contribution of a separate dummy should be interpreted with respect to the reference group.

(c) Changes in other variables and initial income ranks

Finally, term (4) shows that the evolution of IRHI is also determined by the way in which changes in other non-income variables are related to z_{i1} (which is a function of the income rank in the first period). For example, when age is included in x_{it}, term (4) will indicate whether the health effect of getting older ($\beta^{age}(x_{it}^{age} - x_{i1}^{age})$) is related to the initial income rank. In other words, it will isolate the ageing of our cohort from the evolution of IRHI over time.

Summing up, our approach decomposes the evolution of IRHI in four elements and builds on a regression linking health to income and other non-income variables. The first two elements are related to the marginal income distribution only. First, equi-proportionate income growth is associated with rising IRHI when proportional income changes have larger health effects among those with higher incomes. Second, mean-preserving changes in the income distribution lead to rising IRHI when larger health effects from proportional income changes for the rich are combined with pro-rich changes in the income distribution. We re-emphasize that the importance of whether these health effects increase or decrease with income, makes it essential to allow for a flexible functional form of income. The two remaining elements of the decomposition are related to the interdependence between changes in the income distribution and changes in the non-income variables. We find that income mobility (i.e. individuals moving up the income rank) will lead to increasing IRHI when it disproportionally favors those who are already healthy (e.g. the young). We also show that changes in non-income variables only matter insofar as these are related to the initial income rank.

4.4. DATA AND EMPIRICAL IMPLEMENTATION

We use the *China Health and Nutrition Survey* (CHNS) to analyze the evolution of IRHI in China, and the extent to which it is associated with changes in the Chinese income distribution. The CHNS is an ongoing panel data set that covers nine different provinces of China. For details of the CHNS we refer to the project website (Carolina Population Center, 2010).

44. The sum of the associations between the dummy variables and the change in income ranks across all L categories (including the reference category) equals zero since $\sum_{k=1}^{L} x_{it}^{k} = 1$, i.e.

$\sum_{k=1}^{L-1} \sum_i x_{it}^{k} (z_{it} - z_{i1}) = -\sum_i x_{it}^{L} (z_{it} - z_{i1})$.

(a) Estimating the relation between health, income and non-income variables

Our measure of individual health is based on the four possible responses to a self-reported health (SAH) question. In all waves, individuals were asked "Right now, how would you describe your health compared to that of other people of your age?": Excellent, Good, Fair or Poor?[45] These ordered responses are not suitable for measuring inequality with a rank-dependent inequality index such as the Erreygers index since they impose that differences between subsequent SAH categories always represent the same health change (Erreygers and Van Ourti, 2011a). Instead, we use an interval regression estimation method proposed by Van Doorslaer and Jones (2003). It involves the estimation of an ordered probit model with thresholds not estimated but imposed from external data. This interval regression has the advantage of combining (i) the estimation of the relation between income and health in eq 4.2 with (ii) producing a predicted health score that we will use as the measure of health throughout our analyses and that has interval scale properties with a well-defined minimum and maximum value such that it is compatible with the Erreygers index:

$$\text{eq 4.5} \quad h^*_{it} = a + f(y_{it}) + x_{it}'b + e_{it}$$

where h^*_{it} equals latent health, $e_{it} \sim N(0, \sigma^2)$, and a, b, $f(\ )$ are the empirical counterparts of α, β and $\varphi(\ )$ in eq 4.2. So, in eq 4.5 we estimate the association between health and income conditional on a number of x_{it}'s, like described in eq 4.2. However, two types of biases are introduced here. First of all, we estimate eq 4.2 in a sample of Chinese, while eq 4.5 described the whole population. Secondly, we do not observe all relevant non-income variables. This may be problematic, because we estimate unobserved health in our interval regression. If the unobserved variables are correlated with income. Then, the error term may become correlated with income and the association between income and health may be biased. This is an inevitable result, because it follows directly from our chosen approach.

We impose thresholds obtained from the cumulative distribution of the Chinese visual analogue scale (VAS) in the World Health Organization's Multi-Country Survey Study on Health and Responsiveness (Üstün et al., 2003b). A visual analogue scale allows respondents to rate their general health level along a continuous line. The VAS thresholds are obtained by calculating the cumulative frequency of VAS for each SAH category assuming a stable mapping of SAH into VAS (Van Doorslaer and Jones, (2003).

45. If respondents systematically report different levels of SAH for similar objective health states, SAH might be prone to response heterogeneity (Bago d'Uva et al., 2008a). While this will affect the level of IRHI, this is less important for the evolution of IRHI since reporting heterogeneity is most likely largely constant over time in a panel. The information in the CHNS does not allow analyzing this in more detail.

The resulting thresholds are 0.91 (excellent/good), 0.80 (good/fair) and 0.50 (fair/poor); and the minimum and maximum VAS scores are 0 (minimum health) and 1 (maximum health). These thresholds allow us to interpret the predictions from the interval regression as VAS scores. They also show that the assumption of similar health differences between subsequent SAH categories – which would be imposed when using the ordered SAH responses – is unrealistic.

Household income is another crucial variable in this study. The CHNS collects detailed information on various income sources of the household, including income from wages, agriculture, own business and public/private transfers. Household income has been expressed in 2006 Yuan prices (¥100=€9.83) using consumer price indices specific to each wave, province and county; an equivalence scale is used to allow for differences in household size by dividing household income by the square root of household size (Van de Poel et al., 2012). We use a flexible parameterization of the income effect since it is crucial for our decomposition approach to allow for heterogeneity in income effects across poor and rich individuals. We use a polynomial transformation of income as it is differentiable across the entire income range, and let the data determine the order of the polynomial by selecting the most parsimonious polynomial that does not statistically differ from a fifth order polynomial (at a 1 percent significance level).

The other variables x_{it} consist of a set of age-sex dummies, and a set of regional dummies. Sufficiently wide age ranges were adopted (10-29, 30-49, 50-69 and 70+) to ensure that all sex-age categories contain sufficient observations in each period over the 15 year study period. We also account for regional differences by including information on the region of residence. Following Bramall (2009) and Tafreschi (2011), we have divided the CHNS into four regions, defined by coastal/inland province and urban/rural areas reflecting that urban regions, and especially those in coastal provinces, have experienced the strongest economic development in China.

We only include demographic and regional non-income variables in eq 4.5 since we want to provide evidence on the total association between changes in the distribution of income and the evolution of IRHI. We have not included variables that are potentially endogenous and related to the mechanisms underlying this association (such as education, labor force behavior, lifestyles, health insurance) as it would affect the estimate of the magnitude of this association. Hence, we only include variables that are exogenous to the association between income and health. This obviously holds for the standardizing variables age and sex, but in our context also for region since the CHNS does not record migration. For the same reason, we have refrained from using time fixed effects since these could pick up part of the income trend over the 15 year study period.[46] This leaves us with a version of eq 4.5 in which the non-linear income function features as the

46. Adding time and cohort effects to eq 4.5, did not alter our conclusions reached in this paper.

sole potentially endogenous variable. We have deliberately not addressed its potential endogeneity since we are interested in documenting the association between changes in the distribution of income and the evolution of IRHI in China. Turning to the underlying mechanisms is only sensible after the magnitude of this association has been established, and after the relative importance of "income growth", "mean-preserving income changes", and "income mobility" has been understood.

We estimate eq 4.5 on all CHNS waves using a pooled interval regression model, and allow for clustering at the community level (which is the primary sampling unit in the CHNS and therefore leads to conservative statistical inference). It is not feasible to implement an individual fixed effects specification and estimates obtained from a random effects model were very similar to those of the pooled model. Since the random effects model imposes stronger exogeneity assumptions, we prefer the pooled model.

Remember that we highlighted three simplifying assumptions underlying eq 4.2– and thus also eq 4.5. Here, we provide some justification. First, ruling out interactions between the non-linear income function, and age/sex/region is reasonable since alternative versions of the model in eq 4.5 show that most of the interactions are individually insignificant.[47] Our second assumption was to fix the coefficients in eq 4.5 over time. We have checked its validity by adding interactions between time and the non-linear income function, and found almost none of the individual interactions to be statistically significant.[48] Finally, we have ignored dynamics in the specification of eq 4.5 since the period between the various CHNS waves differs from wave to wave.[49]

47. We also derived the counterpart of eq 4.4 that allows for these interactions. This shows that the subdivision in "income growth", "other mean-preserving changes", "income mobility" and "other variables" will change when the magnitude of the age/sex/region effect is strongly affected by the inclusion of these interactions in eq 4.5 for which we found no evidence in the CHNS data.

48. A second sensitivity test consisted of estimating eq 4.5 in a pair-wise manner. In our main analysis in section 4.5, we estimate eq 4.5 on all CHNS waves and use the estimates as an input for each pair-wise decomposition of the change in IRHI between period $_t$ and the first period (using eq 4.4). A straightforward way to relax the assumption of fixed coefficients is to re-estimate eq 4.5 for each pair-wise decomposition between period $_t$ and the first period; and to check whether this affects the results of our decomposition. We found hardly any differences between both approaches, except for the change in IRHI between 2004 and 1991 where we found larger, but still small differences. We take away from this sensitivity test that the assumption of fixed betas is reasonable in the CHNS data.

49. Nevertheless, we have also estimated a version of eq 4.5 that includes a lagged dependent variable (ignoring the fact that the time lags are not constant across waves) and used the resulting estimates as an input for the decomposition. The results confirmed our findings in section 4.5.

(b) Selection of our sample, attrition, descriptives and statistical inference

We use all waves of the CHNS that were available at the time of the analysis (1991, 1993, 1997, 2000, 2004, 2006), except for the first wave (1989) which does not include the SAH question.[50]

Table 4.1 gives an overview of our sample selection criteria. Since our decomposition follows a cohort, we consider only those individuals that were present in 1991 and follow these individuals until 2006 (note that some individuals drop out in a particular year, and reappear in a later wave). We exclude the province of Heilongjiang since it did not participate in the first (1991) wave of our cohort (see row 'Did not participate' in Table 4.1). The province of Liaoning is included, but did not participate in the 1997 wave (see row 'Did not participate', column 1997 in Table 4.1). This will affect our decomposition of the evolution of IRHI between 1991 and 1997, but not the decomposition for any of the other waves. We also delete all respondents under the age of 10 because the CHNS does not report the SAH question for these children in waves 1997, 2000 and 2004. We further drop all CHNS participants that did not report to either the SAH, date of birth, or income

TABLE 4.1: Reasons for dropout and exclusion from CHNS sample

	Wave					
	1991	1993	1997	2000	2004	2006
Exclusion criterion/reason for dropout						
Initial number of observations[a]	27,812	11,577	11,577	11,577	11,577	11,577
Died before interview	0	102	366	569	789	898
Left household before interviewed	1,113[c]	741	0	0	0	0
Did not participate in interview (reason unknown)[e]	11,914	632	3,211	3,423	4,745	3,432
Younger than 10 years	2,540	0	0	0	0	0
Unknown date of birth	8	0	0	0	0	0
No household income	20	19	67	128	61	125
Household income smaller or equal to zero[f]	73	0	0	0	0	0
Household income among highest 2%	252	158	146	147	113	114
No self-reported health	315	96	1,042	2,056	1,217	2,522
Total number of observations	11,577[b]	9,829	6,745	5,254	4,652	4,486

a. Based on unique number of individuals participating in any wave of the CHNS
b. The 11,577 respondents form our cohort at baseline
c. These respondents participated in 1989, but not in 1991
d. Only recorded in 1991 and 1993
e. Includes households that did not participate at all
f. Only applicable in 1991

50. The 2009 wave was not available at the time of the analysis, but – more importantly – does not include the SAH question.

questions. In addition, we remove the two percent highest incomes from each wave[51] and observations with negative incomes in 1991[52] (see rows 'Household income…").

During the 15 years of follow-up, 61.2% of the initial cohort drops out for a variety of reasons such as mortality, moving (out) of the household or refusal to participate in later waves (see Table 4.1). We correct for the impact of non-random dropout by applying inverse probability weights (IPW) to eq 4.4 and eq 4.5 (Jones et al., 2006).[53] We provide further details in the Appendix and Table A. 4.1.

Table 4.2 shows descriptive statistics for our CHNS cohort, both uncorrected and corrected for sample attrition (see below). We observe that health deteriorates with the ageing of our cohort. Average equivalent income more than doubled (in real terms) and this increase was not equally distributed: the Gini index for our cohort grew from 0.330 in 1991 to 0.445 in 2006. Most of the respondents in our initial cohort live in the inland rural areas (45.7%), the cohort shares living in the inland urban and coastal rural areas are approximately equal, while the fewest respondents lived in the coastal urban areas (9.8%).

Finally, we allow in our statistical inference for clustering at the level of the primary sampling units (i.e. communities) and use robust standard errors. This is straightforward for eq 4.5, and for the wave-specific Erreygers indices (O'Donnell et al., 2008), but more difficult for the other elements of our decomposition in eq 4.4 since (i) two consecutive waves of a panel are not independent and because (ii) the combined sampling variability in the estimates of the IPWs, eq 4.5, and proportional income growth (y_t/y_1) should be accounted for. To this end, we bootstrap the entire procedure, i.e. from IPWs to the 4 terms of our decomposition. We draw 2000 bootstrap samples of communities (Mills and Zandvakili, 1997) and address the dependence between panel waves by drawing

51. With the two percent highest incomes included, the marginal effect of income turned negative at very high income levels, presumably due to the bounded nature of health in combination with the flexible, but parameterized polynomial transformation of income. The exclusion of the two percent highest incomes is of little practical importance since the decomposition results obtained from the total sample, including the two percent highest incomes, confirm the results reported in this paper.

52. Negative incomes reflect negative returns of raising livestock, business, farming, gardening and fishing. They are problematic for our decomposition since individuals reporting negative incomes in 1991, will see their incomes decline in the hypothetical scenario of equi-proportional income growth. This seems an unwanted assumption and therefore we exclude all households with incomes smaller or equal to zero Yuan in 1991. Households that report positive incomes in 1991, but negative incomes in later waves are not excluded.

53. Petrie et al. (2011) have shown that IPW corrections for mortality lead to an overestimation of the remaining level of health in the cohort, but this concern is of limited relevance for our analysis since attrition in the CHNS is mainly driven by other causes than mortality (less than 8% of the 1991 cohort have died by 2006, compared to an overall attrition rate of 61.2%). Moreover, application of our decomposition approach to (i) the cohort excluding all individuals that died during the study period, (ii) the balanced and (iii) unbalanced cohorts confirmed the decomposition results reported in this paper.

TABLE 4.2: Descriptives of CHNS 1991 cohort

	Variable	1991	1993	1997	2000	2004	2006
Uncorrected	N	11,577	9,829	6,745	5,254	4,652	4,486
	Percentage females	0.508	0.501	0.491	0.501	0.492	0.499
	Average age	35.7	37.7	42.2	47.2	51.6	53.4
	Region						
	Coastal urban	9.8%	9.1%	4.9%	8.0%	6.2%	6.6%
	Coastal rural	24.0%	25.2%	19.9%	25.6%	25.5%	25.5%
	Inland urban	20.5%	19.2%	21.3%	19.7%	18.7%	17.9%
	Inland rural	45.7%	46.5%	54.0%	46.6%	49.6%	50.1%
	Income*	¥4,823	¥5,354	¥6,508	¥7,756	¥8,833	¥10,270
	Income distribution (gini)	0.330	0.370	0.371	0.403	0.432	0.443
	Income percentiles (5-95)	(1,012-10,452)	(914-12,783)	(985-15,976)	(705-19,750)	(896-24,246)	(747-27,770)
	SAH response						
	Excellent	13.6%	12.6%	11.3%	11.8%	10.5%	8.6%
	Good	61.7%	63.6%	60.9%	47.9%	42.6%	44.2%
	Fair	21.1%	20.1%	23.4%	33.2%	37.8%	38.0%
	Poor	3.7%	3.8%	4.5%	7.1%	9.1%	9.2%
	Average health	0.812	0.809	0.800	0.792	0.780	0.777

TABLE 4.2: Descriptives of CHNS 1991 cohort (continued)

	Variable	1991	1993	1997	2000	2004	2006
Corrected for dropout	Percentage females	0.508	0.508	0.512	0.51	0.514	0.51
	Average age	35.7	37.5	41.4	45.2	49.8	51.8
	Region						
	Coastal urban	9.8%	9.9%	10.7%	10.7%	12.2%	11.4%
	Coastal rural	24.0%	24.0%	24.3%	24.1%	23.8%	23.4%
	Inland urban	20.5%	20.5%	20.3%	21.0%	21.4%	22.5%
	Inland rural	45.7%	45.6%	44.7%	44.3%	42.6%	42.7%
	Income*	¥4,823	¥5,424	¥6,817	¥7,861	¥9,556	¥11,089
	Income distribution (gini)	0.330	0.369	0.371	0.401	0.430	0.445
	Income percentiles (5-95)	(1,002-10,452)	(933-12,965)	(984-16,600)	(709-20,018)	(918-26,177)	(617-29,229)
	SAH response						
	Excellent	13.6%	12.7%	13.0%	13.0%	12.5%	9.9%
	Good	61.7%	63.3%	61.1%	48.4%	44.4%	46.2%
	Fair	21.1%	20.1%	22.6%	32.0%	37.0%	37.9%
	Poor	3.7%	3.9%	4.4%	7.0%	8.5%	8.2%
	Average health	0.812	0.809	0.803	0.796	0.787	0.784

Note: Income is average equivalent household income (¥, 2006)

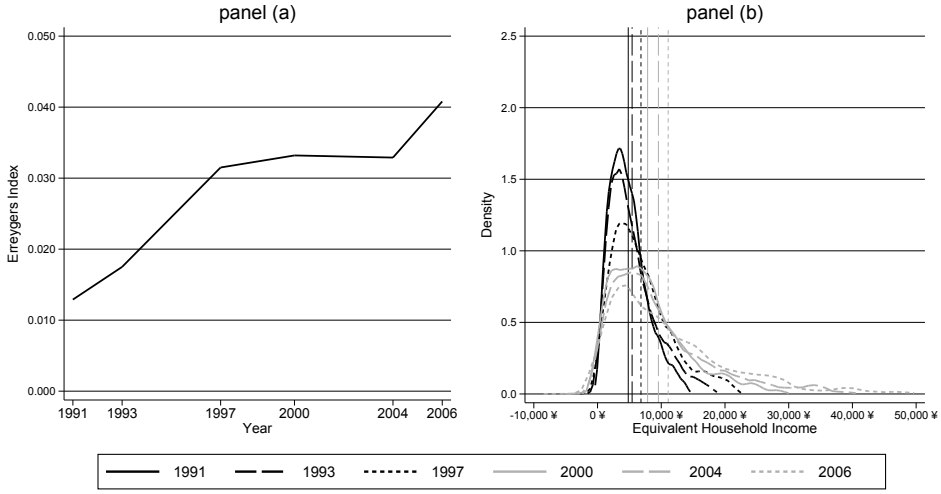
bootstrap samples of the communities in the first wave only (and use these same communities for all later waves).

4.5. RESULTS

Table 4.2 showed that the 1991 cohort of the CHNS experienced a dramatic increase in average incomes and income inequality over the period 1991-2006. In this section, we link these changes to the evolution of IRHI.

(a) Evolution of IRHI and changes in the income distribution in China

The trends in Table 4.2 are illustrated in greater detail in panel b of Figure 4.1 showing that, on top of proportional income growth, the spread of incomes has increased between 1991 and 2006. Panel a of the same figure shows that IRHI more than doubled from 0.013 in 1991 to 0.041 in 2006, and this trend has favored the rich throughout as indicated by the positive values of the Erreygers indices. Our finding of a large increase in IRHI between 1991 and 2006 is a first and important finding of this paper, and the starting point for the subsequent analyses.



Note: the vertical bars in panel (b) represent the average equivalent income levels per wave; results are obtained after applying inverse probability weights (IPW) to eq 4.4 and eq 4.5.

FIGURE 4.1: Evolution of IRHI and changes in the density function of equivalent household income in China from 1991-2006

(b) The relation between health, income and non-income variables in China

An essential ingredient to understand these rising IRHI is the relationship between health and income conditional on the non-income variables. The interval regression estimates of

TABLE 4.3: Health equation estimates and drivers of differential income mobility

Variable	Interval regression coefficient[a,b]	share of individuals in age-sex / region category k					difference between average weights z_{it}^{c}				
		1993	1997	2000	2004	2006	1993	1997	2000	2004	2006
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Females											
10_29	Reference	0.194	0.147	0.069	0.031	0.021	32	619	1,021	2,028	2,445
30_49	-0.041 ***	0.183	0.206	0.262	0.258	0.245	112	194	181	86	346
50_69	-0.105 ***	0.100	0.113	0.126	0.148	0.157	-287	-910	-608	-565	-544
70_89	-0.149 ***	0.030	0.045	0.054	0.077	0.087	-86	-356	-1,168	-695	-1,575
Males											
10_29	0.005 **	0.192	0.153	0.104	0.067	0.040	76	373	338	814	1,547
30_49	-0.022 ***	0.177	0.185	0.211	0.200	0.213	64	135	109	-34	129
50_69	-0.084 ***	0.100	0.112	0.127	0.149	0.162	-215	-466	-173	-108	-165
70_89	-0.138 ***	0.022	0.038	0.048	0.069	0.076	-44	-1,208	-1,015	-292	-202
Region											
coastal – urban	0.010 **	0.099	0.107	0.107	0.122	0.114	-38	-304	-121	634	5
coastal – rural	0.023 ***	0.240	0.243	0.241	0.238	0.234	111	482	571	323	191
inland – urban	-0.006 ***	0.205	0.203	0.210	0.214	0.225	-123	-915	-820	-1,207	-880
inland - rural	Reference	0.456	0.447	0.443	0.426	0.427	-1	172	24	28	150
Equivalent income/1,000											
linear	3.442E-03 ***										
quadratic	-1.486E-04 **										
cubic	2.247E-06 **										
Constant	0.832 ***										

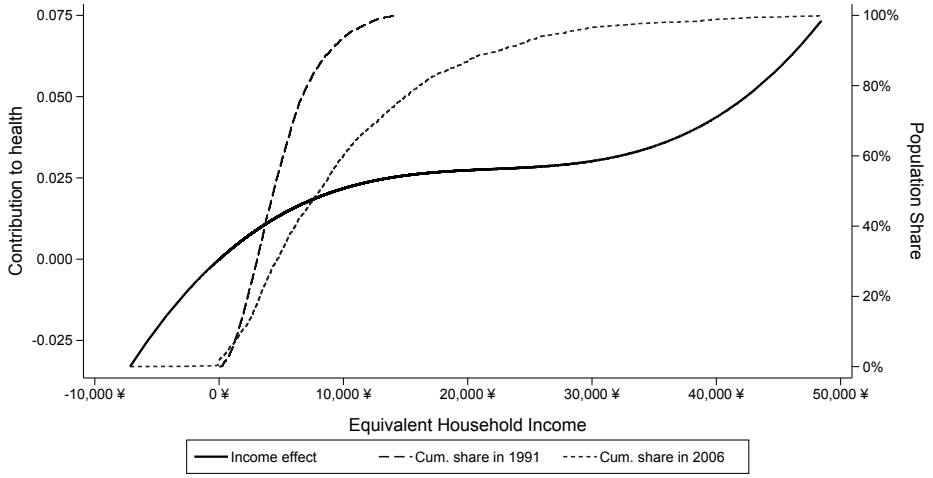
Note: Results are obtained after applying inverse probability weights (IPW) to eq 4.4 and eq 4.5.

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

a. Number of observations in pooled interval regions is 42,542; the division over the waves can be found in Table 4.2

b. Model statistics: Wald chi(13) = 2,712 (P<0.0001) LR cubic/quadratic income=23.48 (P<0.0001)

c. $(z_{it}/n_t^k) - (z_{i1}/n_1^k)$ equals the difference between the average weight z_{it} for age-sex/region category k in periods t and 1.



Note: results are obtained after applying inverse probability weights (IPW) to eq 4.4 and eq 4.5. Decomposition of evolution vof IRHI in China

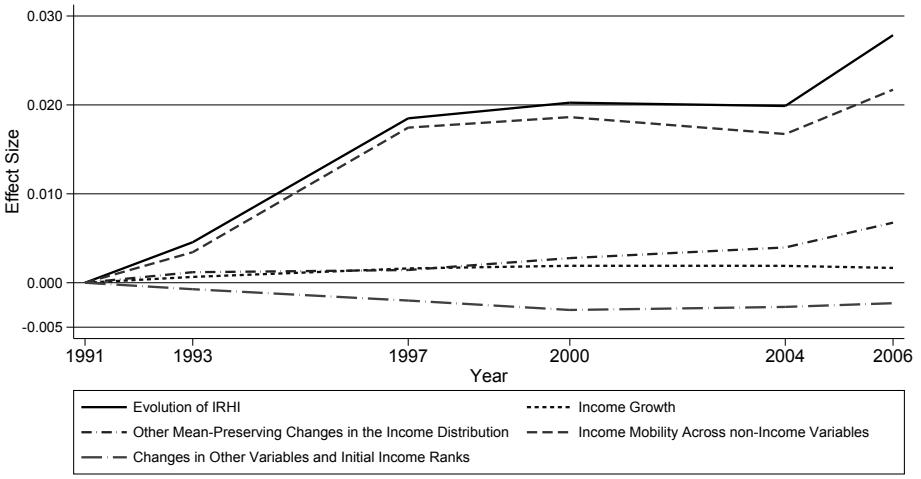
FIGURE 4.2: Effect of income on health and cumulative populations shares in first and final wave by income

4

eq 4.5 reported in column (1) of Table 4.3 show that demographics play a major role: (VAS-scaled) SAH deteriorates monotonically with age and is higher for males than females. We also find that those living in coastal and rural areas report better health, which is consistent with Van de Poel et al. (2012) who found urbanization in China to raise the probability of reporting poor health. Finally, the relationship between income and health is best described by a third degree income polynomial, giving rise to the income profile in Figure 4.2: health increases with income at a decreasing rate, but for very high incomes (~ inflection point at 22,048 Yuan), the marginal effect of additional income starts rising again. The cumulative income distributions for 1991 and 2006 in Figure 4.2 confirm that this only occurs for a minor share of our cohort: in 1991, *no* individuals had an income level above the point of inflection, whereas in 2006, this was about 13% percent.

While this cannot be inferred from Figure 4.2, the underlying estimates in Table 4.3 indicate that proportional income changes have overall similar health effects among those with lower, middle and higher incomes (see also section 4.3(b)). It follows that according to our decomposition framework, proportional income growth will – *ceteris paribus* – be associated with constant IRHI. This is a second, and perhaps more surprising, finding of our paper: despite the enormous improvement of average incomes in China over the last few decades, these proportional income gains do not help us understand the rising IRHI in China.

Figure 4.3 and Table A. 4.2 in the appendix show estimates of the 4 terms of our decomposition: (i) income growth; (ii) other mean-preserving changes in the income distribution; (iii) differences in income mobility across non-income variables; and (iv)



Note: statistical inference of the decomposition is provided in Table A. 4.2; results are obtained after applying inverse probability weights (IPW) to eq 4.4 and eq 4.5.

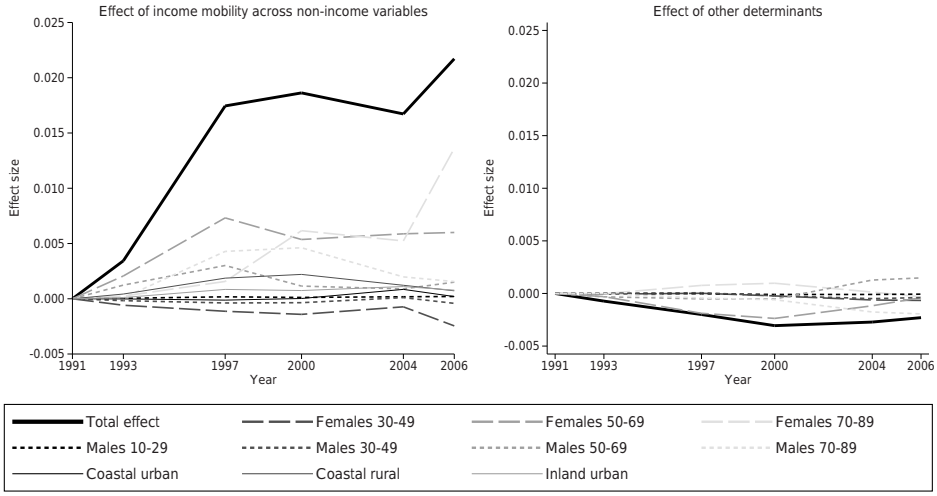
FIGURE 4.3: Decomposition of the evolution in IRHI between year t and 1991

differences in the evolution of non-income variables across the initial income distribution. We discuss each of these terms, including statistical inference, in detail.

First, proportional income growth only shows a minor (and statistically insignificant) association with IRHI. This means that income growth in China is not related to rising IRHI, as was already anticipated by the earlier finding that proportional income changes have an overall uniform health effect across the income distribution.

Second, compared to proportional income growth, 'other mean-preserving changes in the income distribution' are more important for IRHI. Given that proportional income increases are associated to a similar health gains across the income spectrum, this must be due to the pro-rich changes in the Chinese income distribution that we reported in Figure 4.1 and Table 4.2. Hence, our results confirm that the rise in Chinese income inequality is associated with increasing IRHI, but it also confirms that the other elements of our decomposition must be more important: in none of the waves the 'mean-preserving changes' account for more than 25 percent of the change in IRHI (and in several waves far less). It is also worth noting that the 'direct' effects of changes in the income distribution (i.e. term (1) and term (2)) only account for around 30 percent of the rise in IRHI.

Third, differences in the evolution of the non-income variables (term 4) show a small, but statistically significant, negative effect on IRHI in Figure 4.3. This effect is entirely driven by the evolution of age since sex and region are time-invariant in the CHNS. In other words, this term isolates the impact of ageing on the change of IRHI over time. We find that ageing contributes to IRHI (i.e. IRHI would have increased slightly more



Note: statistical inference is provided in Table A. 4.2; results are obtained after applying inverse probability weights (IPW) to eq 4.4 and eq 4.5.

FIGURE 4.4: Differences in income mobility across age, sex and region (panel a) and differences in non-income variables across initial income (panel b)

without ageing) due to the initially rich experiencing larger health declines due to ageing compared to the initially poor. This is confirmed by the right panel of Figure 4.4 which shows the separate contributions of the different age-sex groups.[54]

It is clear by now that most of the change in IRHI must be related to differential income mobility across the non-income variables (term 3 in eq 4.4). Differential income mobility by age, sex and region accounts for around 70 to 90% of the total evolution in IRHI (see Figure 4.3). This means that individuals that are sliding downward in the income distribution are less healthy compared to the reference group of young (and healthy) females living in inland-rural areas. The left panel of Figure 4.4 and Table A. 4.2 illustrate that regional differences in income mobility have a small, and insignificant effect (reference category: inland-rural), and they are clearly dominated by the differential income mobility across sex-age categories (youngest females are the reference). There is no statistically significant difference between males younger than 50 and females younger than 30 (the reference category), and for females in the age group 30-50, we observe a negative

54. For example, we find a relatively large negative effect for females in age group 50-69. As an example, we explain how this negative effect should be interpreted; a similar reasoning applies to the other age-sex groups. First, the female 50-69 age group has a large negative coefficient β^k in the interval regression (see also column 1 in table 3). Consequently, for a female to contribute negatively to the 50-69 age group, it must be that $(x_{it}^k - x_{i1}^k)\, z_{i1} > 0$. This will happen when she enters the 50-69 group and is relatively rich, i.e. $(x_{it}^k - x_{i1}^k) > 0$ and $z_{i1} > 0$, or when she leaves this age group and is rather poor, i.e. $(x_{it}^k - x_{i1}^k) < 0$ and $z_{i1} < 0$. Second, females not changing age groups, do not contribute to term 4.

insignificant effect. By far the strongest effects are observed for males and females above 50 years who were downwardly income mobile and less healthy than the younger age groups. The effects for 50+ females are in most cases also substantially larger and more significant than for 50+ males.

This begs the question why the effect for elderly females is that much larger than for elderly males. Closer inspection of term 3 in eq 4.4 reveals that the effect for a given age-sex category is driven by three components: (i) the health of this age-sex category (compared to the reference group), (ii) the share of people in this category, and/or (iii) whether this group has been upwardly or downwardly income mobile.[55] In other words, elderly females can show a larger effect than elderly males for three reasons: because they report lower health, because there are more elderly women in our cohort in any given wave, or because elderly women are more downwardly mobile than elderly males. Table 4.3 presents some evidence on the relative importance of these three elements for all age-sex (and regional) categories. It shows that the diverging patterns between elderly women and men in Figure 4.4 are mainly related to differences in income mobility (see columns 7-11), and less to the lower health level of females (column 1) or their relative group size (columns 2-6), although these clearly also matter.

We thus conclude that differential income mobility across age-sex groups is the single most essential element to understand the evolution of IRHI in China. Young (and healthier) individuals have been much more upwardly mobile compared to those above 50, and this age difference is more pronounced among females than males, in particular among those older than 70. While the inferior health of older women and their increasing group size over time have also contributed to this phenomenon, downward income mobility of elderly women (compared to young women) is much more important. We have also established that rising income inequalities account for less than a quarter of the change in IRHI and that ageing in China even slightly reduces IRHI. Perhaps the most surprising result is the lack of any association between income growth and rising IRHI in China.

(c) Further unraveling the downward income mobility of elderly (females): an exploration

Despite its importance for understanding the evolution of IRHI in China, Table 4.3 does not explain why the 50+ (and in particular women in the 70+ group) have experienced

55. Since we use IPWs, an additional source of variation emerges, i.e. the difference between the IPW-weighted age-sex (and regional) composition and the one that would have prevailed without attrition. Only in the unlikely case when the IPWs predict attrition perfectly, this source of variation will vanish. We do indeed find that the IPWs are not perfect, but its effect on term 3 of our decomposition is negligible compared to the other three drivers, and if anything, reduces – rather than raises – the effect of the 70+. In order not to complicate the results too much, we do not report this additional source of bias. More information can be obtained from the authors.

a greater downward shift in their income positions since 1991 compared to younger and healthier generations. In this section we explore a potential explanation.

A first useful observation is that our analysis has been based on equivalent incomes (see also section 4.4(a)). Hence, the downward shifts can be driven by changes in either the numerator (household income) or the denominator (household size), or both. It turns out that the diverging pattern for the elderly compared to younger age cohorts – and for elderly women compared to men – is not driven by changes in household size, but by changes in household incomes.[56] Therefore, explanations based on changes in household composition *only* cannot explain differential income mobility between the young/healthy and the old/unhealthy (and amongst the elderly) in China.

Instead, an explanation that focuses on changes in household incomes squares with evidence on replacement incomes and income support for the elderly in China. Elderly in rural areas – especially those above 70 – typically do not receive any pension benefits and still largely rely on family-based self-insurance mechanisms. These insurance mechanisms have become more uncertain due to a combination of China's demographic transition leading to smaller family sizes, less co-residence and fewer extended families (Zhong, 2011)[57] as well as to migration of (mainly young) adults from rural to urban areas (Giles et al., 2010). Similarly, urban dwellers – in spite of having more favorable pension arrangements than rural citizens – have been confronted with pension arrears, due to bankruptcies of state-owned enterprises and due to limited compliance with recent pay-as-you-go local pension initiatives. Traditional family-based insurance mechanisms have only partially protected elderly urban residents from these pension arrears (Cai et al., 2006). Nevertheless, urban residents have 'on average' continued to receive pension benefits that are largely unavailable to rural citizens (Feldstein, 1999; Yang, 1999). Giles et al. (2010) show that less than 5 percent of elderly rural residents relied on pension incomes in 2005 whereas this share was 45 percent in urban areas. Moreover, and in line with our findings in Table 4.3, these authors report that pension incomes are more concentrated among elderly males in both rural and urban areas, while females rely more on family-based income support which, according to Meng et al. (2007), is explained by higher former labor market participation of males. Finally, Knight et al. (2004) show that rising income inequality in Chinese coastal cities was mainly related to changes in

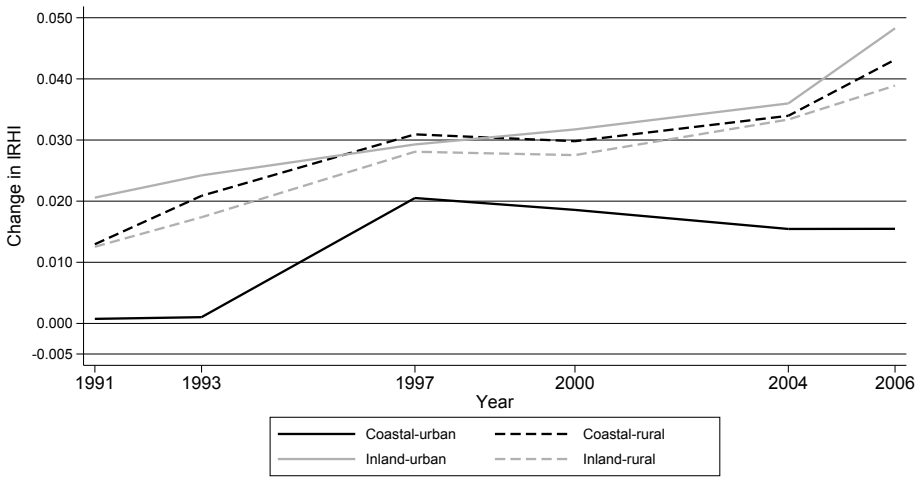
56. For example, the average female (male) in the 10-29 age range in 2006 had seen her (his) household income rise with a ratio of 3.04 (3.09) while this ratio was only 1.25 (1.83) for the average female (male) in the oldest age group. Similar ratios for household size amounted to 1.07 and 0.83 for females and 1.09 and 0.75 for men.

57. Zhong (2011) also finds that population ageing accounts for a substantial share of rising income inequality in rural China, but not in urban China.

wage incomes whereas pension incomes accounted for the largest share of rising income inequalities in inland urban areas.

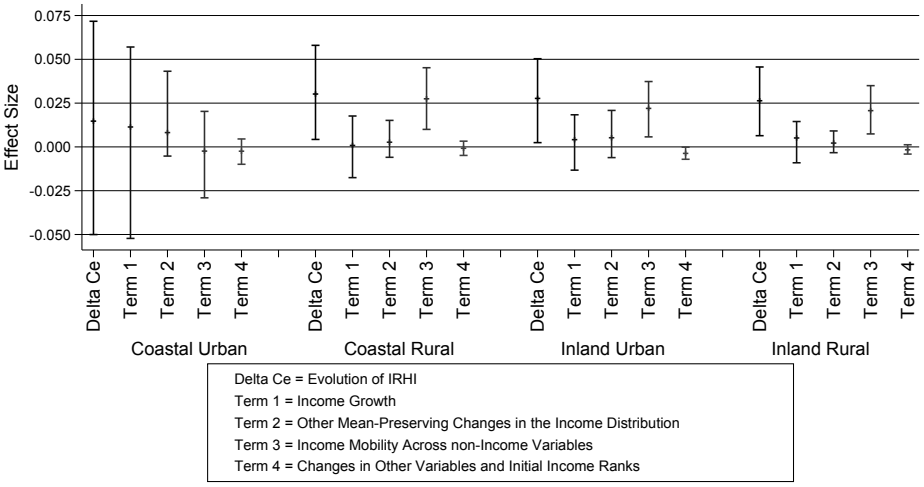
Two lessons emerge from this brief review for this paper. First, it confirms that our findings of differential income mobility between the young and old (and between elderly males and females) are in line with the evidence to date. Secondly, and more importantly, it points to a potentially crucial difference between rural and urban areas and between inland-urban and coastal-urban areas. All of these studies suggest that differential income mobility should play a more important role in rural and inland-urban areas than in coastal urban areas. Hence, we repeated our decomposition of IRHI in section 4.5(c) (and the interval regression models in section 4.5(b)) separately for the four regions defined by inland/coastal and urban/rural. While this leaves us with more than 1000 observations for each of the rural regions, the number of observations in the urban regions becomes fairly small in the last wave of the CHNS, i.e. respectively 294 and 802 respondents in the urban-coastal and urban-inland areas.

Figure 4.5 shows the evolution of IRHI for the four different regions. Inequalities were highest in the two inland and coastal-rural areas, and the magnitude of IRHI in these regions was comparable to that of the entire sample (in Figure 4.1). The urban areas in the coastal provinces followed a different pattern and IRHI were smaller (and insignificant) in all periods. For each region, we present the decomposition of the evolution of IRHI between 1991 and 2006 in Figure 4.6. These decompositions show a similar grouping for the two inland and coastal-rural areas versus the coastal urban area. The decomposition of the former three areas is very similar to that of the entire sample (but significance



Note: results are obtained after applying inverse probability weights (IPW) to eq 4.4 and eq 4.5. Statistical inference is described in the text. ore details can be obtained from the authors

FIGURE 4.5: Evolution of IRHI in coastal/inland and urban/rural areas



Note: results are obtained after applying inverse probability weights (IPW) to eq 4.4 and eq 4.5.

FIGURE 4.6: Decomposition of IRHI in coastal/inland and urban/rural areas in 2006 with 99% CI

levels are lower) and is also here dominated by differences in income mobility across age groups. In contrast, income growth and other mean-preserving changes in the income distribution are more substantive in the coastal urban areas, while differential income mobility across age-sex categories is negligible. However, none of the terms of the decomposition for the urban-coastal region is still significant – probably as a result of small sample size – and we thus have to be careful in interpreting these as contributions to the evolution of IRHI.

Our findings per region thus suggest that while overall in China IRHI are to a large extent driven by differences in income mobility between the young and the elderly, and between sexes amongst the elderly, regional differences in income protection (via family based insurance mechanisms or more formal pension systems) appear crucial to understand this phenomenon. It also means that the younger cohorts have gained most from the economic prosperity in recent decades in China, except for the coastal urban areas where all age groups seem to have benefited equally.

4.6. CONCLUSIONS

The rapid economic growth in China in recent decades has been accompanied by larger inequalities in society, not only with respect to income, but also with respect to health. This begs the question how the dramatic changes in both the level and distribution of Chinese household incomes over the last few decades have been related to the extent to which also health – as self-reported – is distributed unequally by income. We believe we make the following contributions. First, we propose a new cohort-decomposition of

the changes in IRHI into four components: (i) the growth of mean income, (ii) mean-preserving changes in the distribution of incomes, (iii) the income mobility across other important variables like age and region, and (iv) changes in initial income rank. It differs from earlier approaches in two respects: (a) it extends and simplifies earlier decomposition approaches by replacing the standard concentration index by the Erreygers index which recognizes that health – unlike income – has a finite maximum; (b) it highlights the importance of income mobility when it varies across important determinants of health like age and region. Secondly, we empirically implement the approach using data from the CHNS, one of the longest running panel studies, by estimating health as a polynomial function of income, controlling for age, gender and region. While this does not allow for causal inference, the decomposition analysis does reveal which factors are important – and which are not – for understanding the evolution of health disparities by income in China.

Our findings are as follows. We find that IRHI rose quite substantially in this cohort: between 1991 and 2006 its degree has more than tripled. No other study had revealed this rise before, and in the absence of ageing of our cohort the rise would have been larger. Also new is that this substantive rise turns out to be largely unrelated to the double digit average income growth rates and only weakly related to the overall growth in income inequality. Instead, what emerges is that most of the increase in IRHI relates to the very different income mobility by age and gender. Especially women over 50 in the cohort have experienced much more adverse trends than other groups in their income position (and their health) and this explains more than half of the growth in IRHI. This strongly suggests that elderly women were left behind in the rapid economic developments occurring in China and that low or absent replacement incomes for sick and old women are a key factor here. While we have not directly tested this hypothesis, it conforms to the literature on retirement incomes. This confirms that, especially in rural areas, pension entitlements are largely lacking for the majority of the population. Reliance on traditional, family-based self-insurance has also declined because of demographic transitions like smaller family sizes and migration. Urban and male residents have generally enjoyed more favorable pension arrangements because of their greater participation in the formal labor market. This explanation also squares with the regional differences in our findings: income mobility did not contribute as much to IRHI changes in the coastal urban areas as in the other regions. This suggests that income protection and pension arrangements, also for older women, in this region are better than in other parts of China.

Our study is thus a first step in understanding the evolution of IRHI in China and its relation with changes in the income distribution. Our findings strongly suggest that the following issues deserve attention in further research. First, how are age profiles of income related to age profiles of health in China, and in particular why are low replace-

ment incomes associated with low health? Our results suggest that this is a structural relationship and Giles et al. (2010) emphasize that many elderly without any replacement income are still working. It is most likely that this has consequences for their health. Secondly, it appears relevant to investigate further the different experiences in urban coastal areas with other regions to find out which differences in institutions versus demographics matter. Finally, it would be of interest to examine whether differential income mobility among the elderly is related to changes in household composition due to mortality or migration of income earning members.

APPENDIX

Implementation of inverse probability weights

The IPW-methodology consists of two distinct steps. First, for all respondents in the first wave, one estimates the probability of still being present in wave t. Second, the inverse of the resulting predicted probabilities are used as probability weights in wave t, i.e. as an IPW, in all analyses.

For each wave, we estimate a separate probit, and include the following wave 1 regressors: (a) a polynomial of equivalent income, (b) dummies for the separate SAH categories, and (c) sex-age dummies and regional dummies. The results in Table A. 4.1 show that the middle aged (30-70) are most likely to stay in the cohort, and that young females are most likely to leave the cohort (in later waves elderly males are least likely to stay probably due to mortality). We also observe that those living in rural areas[58] and those in good health are most likely to stay in the cohort. The effect of income is non-linear in most waves and indicates that those with higher incomes (except for those with very low and very high incomes) are more likely to drop out of the cohort. Most importantly, we found that the interactions between the income polynomial and the set of SAH dummies were jointly insignificant and therefore removed these from the probit models. This indicates that selective health-income dropout is unlikely to be a serious problem in our data.

While the IPWs should in principle be used in all analyses, we only use them for the estimation of eq 4.5, for the weighted income z_{it}'s[59] (Lerman and Yitzhaki, 1989), for calculating average income in each wave y_t which is needed to calculate y_{it}^{pg}, and for the

58. Note that the coefficients for the coastal regions are much more negative in 1997 since the province Liaoning did not participate in the CHNS in that wave.

59. $z_{jt}^{IPW} = 0.5\left(\sum_{l=1}^{n_t} IPW_{lt}\right) - \left(\sum_{l=1}^{j-1} IPW_{lt} + 0.5\, IPW_{jt}\right)$

left-hand side Erreygers indices in eq 4.4.[60] The IPW's are partially needed for terms (2) and (3) and not for terms (1) and (4) of eq 4.4. The latter two terms can be calculated without the IPWs since we know what would have been the age, sex, and region of individuals that have dropped out of the cohort. Similarly, we can calculate the incomes in the 'proportional income growth' scenario since equivalent income in the first period is known.

TABLE A. 4.1: Probit models of the probability to be in wave t conditional on wave 1 characteristics

	1993		1997		2000		2004		2006	
Variable	Coef	P-value	Coef	P-value	Coef	P-value	Coef	P-value	Coef	P-value
Females										
10_29	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
30_49	0.674	0.000	0.739	0.000	1.118	0.000	1.274	0.000	1.226	0.000
50_69	0.422	0.000	0.552	0.000	0.723	0.000	0.857	0.000	0.791	0.000
70_89	-0.030	0.730	-0.008	0.921	-0.055	0.520	-0.226	0.020	-0.439	0.000
Males										
10_29	0.250	0.000	0.282	0.000	0.255	0.000	0.331	0.000	0.276	0.000
30_49	0.558	0.000	0.680	0.000	0.907	0.000	1.117	0.000	1.086	0.000
50_69	0.380	0.000	0.455	0.000	0.644	0.000	0.759	0.000	0.633	0.000
70_89	-0.110	0.261	-0.113	0.223	-0.218	0.036	-0.465	0.000	-1.100	0.000
Region										
Coastal-urban	-0.268	0.000	-1.049	0.000	-0.277	0.000	-0.516	0.000	-0.500	0.000
Coastal-rural	0.129	0.001	-0.584	0.000	0.002	0.958	-0.073	0.021	-0.100	0.002
Inland-urban	-0.233	0.000	-0.217	0.000	-0.086	0.010	-0.179	0.000	-0.247	0.000
Inland-rural	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
Income										
linear	4.852E-01	0.002	1.757E-01	0.017	-1.862E-02	0.000	8.328E-02	0.267	-9.428E-02	0.000
quadratic	-2.174E-01	0.001	-7.389E-02	0.001			-5.136E-02	0.022	5.694E-03	0.000
cubic	3.758E-02	0.003	8.920E-03	0.000			6.524E-03	0.010		
power four	-2.870E-03	0.005	-3.320E-04	0.000			-2.467E-04	9.739E-03		
power five	8.015E-05	0.008								
SAH response										
Excellent	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
Good	0.019	0.673	0.034	0.358	0.003	0.939	0.052	0.172	0.028	0.466
Fair	-0.019	0.724	-0.003	0.946	-0.025	0.563	-0.008	0.852	-0.026	0.560
Poor	-0.418	0.000	-0.420	0.000	-0.420	0.000	-0.420	0.000	-0.383	0.000
Constant	0.618	0.000	0.124	0.156	-0.486	0.000	-0.656	0.000	-0.477	0.000

60. $C_E^{IPW}(h_t|\ y) = \left[8 \Big/ \left(\sum_{j=1}^{n_t} IPW_{jt}\right)^2\right] \sum_{j=1}^{n_t} \left(IPW_{jt}\, z_{jt}\, h_{jt}\right)$

TABLE A. 4.2: Decomposition Results

Variable	1991	1993	1997	2000	2004	2006
N	11,577	9,829	6,745	5,254	4,652	4,486
Average Health	0.8119	0.809	0.8031	0.7957	0.7868	0.7837
Erreygers	0.0130	0.0175	0.0315	0.0332	0.0329	0.0408
	(0.0049; 0.0222)	(0.0082; 0.0260)	(0.0211; 0.0402)	(0.0237; 0.0415)	(0.0169; 0.0459)	(0.0222; 0.0556)
IRHI change w.r.t. 1991		0.0046	0.0185	0.0203	0.0199	0.0278
		(-0.0011; 0.0086)	(0.0105; 0.0252)	(0.0114; 0.0277)	(0.0053; 0.0331)	(0.0100; 0.0432)
Main Effects						
Other Mean preserving Changes in the Income Distribution		0.0012	0.0014	0.0028	0.0040	0.0068
		(0.0004; 0.0024)	(0.0005; 0.0039)	(0.0005; 0.0077)	(-0.0018; 0.0159)	(0.0005; 0.0158)
Income Mobility Across non-Income Variables		0.0034	0.0174	0.0186	0.0167	0.0217
		(-0.0015; 0.0078)	(0.0111; 0.0245)	(0.0111; 0.0257)	(0.0085; 0.0249)	(0.0090; 0.0358)
Income Growth		0.0007	0.0016	0.0019	0.0019	0.0017
		(-0.0011; 0.0014)	(-0.0036; 0.0041)	(-0.0049; 0.0058)	(-0.0077; 0.0091)	(-0.0113; 0.0132)
Changes in Other Variables and Initial Income Ranks		-0.0007	-0.0020	-0.0031	-0.0027	-0.0023
		(-0.0015; 0.0000)	(-0.0032; -0.0007)	(-0.0047; -0.0016)	(-0.0045; -0.0009)	(-0.0039; -0.0007)
Contributions of Individual Variables to Income Mobility Across non-Income Variables						
Females						
10_29		Reference	Reference	Reference	Reference	Reference
30_49		-0.0006	-0.0011	-0.0014	-0.0007	-0.0025
		(-0.0014; 0.0002)	(-0.0025; 0.0002)	(-0.0033; 0.0003)	(-0.0030; 0.0016)	(-0.0052; 0.0006)
50_69		0.0021	0.0073	0.0054	0.0059	0.0060
		(-0.0003; 0.0041)	(0.0047; 0.0101)	(0.0025; 0.0081)	(0.0030; 0.0090)	(0.0024; 0.0101)

TABLE A. 4.2: Decomposition Results (cont.)

Variable	1991	1993	1997	2000	2004	2006
70_89		0.0002	0.0016	0.0062	0.0052	0.0136
		(-0.0014; 0.0019)	(-0.0011; 0.0044)	(0.0023; 0.0105)	(0.0002; 0.0100)	(0.0072; 0.0213)
Males						
10_29		0.0000	0.0002	0.0001	0.0002	0.0002
		(-0.0001; 0.0002)	(0.0000; 0.0005)	(0.0000; 0.0004)	(0.0000; 0.0005)	(0.0000; 0.0006)
30_49		-0.0002	-0.0004	-0.0004	0.0001	-0.0004
		(-0.0007; 0.0004)	(-0.0012; 0.0002)	(-0.0012; 0.0003)	(-0.0008; 0.0010)	(-0.0018; 0.0010)
50_69		0.0012	0.0030	0.0011	0.0009	0.0015
		(-0.0004; 0.0027)	(0.0009; 0.0054)	(-0.0009; 0.0035)	(-0.0015; 0.0034)	(-0.0016; 0.0051)
70_89		0.0001	0.0043	0.0046	0.0020	0.0016
		(-0.0012; 0.0014)	(0.0009; 0.0092)	(0.0007; 0.0084)	(-0.0071; 0.0087)	(-0.0206; 0.0202)
Region						
inland - rural		Reference	Reference	Reference	Reference	Reference
inland – urban		0.0001	0.0008	0.0007	0.0011	0.0008
		(-0.0006; 0.0009)	(-0.0007; 0.0030)	(-0.0006; 0.0033)	(-0.0009; 0.0037)	(-0.0006; 0.0033)
coastal – rural		0.0004	0.0019	0.0022	0.0012	0.0007
		(-0.0020; 0.0030)	(-0.0008; 0.0049)	(-0.0002; 0.0054)	(-0.0018; 0.0045)	(-0.0027; 0.0038)
coastal – urban		0.0000	-0.0001	0.0000	0.0008	0.0002
		(-0.0011; 0.0006)	(-0.0034; 0.0007)	(-0.0022; 0.0010)	(-0.0012; 0.0031)	(-0.0020; 0.0032)

TABLE A. 4.2: Decomposition Results (cont.)

Variable	1991	1993	1997	2000	2004	2006
Contributions of Individual Variables to Changes in Other Variables and Initial Income Ranks						
Females						
10_29		Reference	Reference	Reference	Reference	Reference
30_49		-0.0001	0.0000	-0.0002	-0.0006	-0.0007
		(-0.0005; 0.0004)	(-0.0008; 0.0009)	(-0.0013; 0.0008)	(-0.0021; 0.0007)	(-0.0021; 0.0006)
50_69		-0.0003	-0.0019	-0.0024	-0.0012	-0.0004
		(-0.0011; 0.0004)	(-0.0034; -0.0004)	(-0.0046; -0.0003)	(-0.0040; 0.0016)	(-0.0036; 0.0026)
70_89		0.0000	0.0008	0.0010	0.0001	-0.0003
		(-0.0005; 0.0006)	(-0.0005; 0.0021)	(-0.0007; 0.0026)	(-0.0022; 0.0024)	(-0.0030; 0.0022)
Males						
10_29		0.0000	0.0000	-0.0001	-0.0001	-0.0001
		(-0.0001; 0.0000)	(-0.0001; 0.0001)	(-0.0003; 0.0000)	(-0.0003; 0.0000)	(-0.0002; 0.0000)
30_49		0.0000	0.0001	-0.0002	-0.0005	-0.0004
		(-0.0002; 0.0003)	(-0.0004; 0.0006)	(-0.0009; 0.0004)	(-0.0012; 0.0002)	(-0.0012; 0.0004)
50_69		-0.0003	-0.0005	-0.0005	0.0013	0.0015
		(-0.0011; 0.0004)	(-0.0020; 0.0009)	(-0.0025; 0.0013)	(-0.0011; 0.0036)	(-0.0011; 0.0041)
70_89		0.0000	-0.0004	-0.0006	-0.0018	-0.0019
		(-0.0006; 0.0004)	(-0.0017; 0.0010)	(-0.0024; 0.0012)	(-0.0040; 0.0004)	(-0.0046; 0.0006)

Note: 99% confidence intervals between brackets; results are obtained after applying inverse probability weights (IPW) to eq 4.4 and eq 4.5.

Chapter 5

Healthy life expectancy around the globe using the World Health Surveys: Illustration and methodological considerations

In collaboration with: Teresa Bago d'Uva

ABSTRACT

Background

Promoting health is the defining goal of health systems and international comparisons can help evaluate performance. The work of the World Health Organization's (WHO) has been at the forefront of the development of population health measures like healthy life expectancy. In this study we propose two modifications to the WHO methodology and provide new empirical estimates of healthy life expectancy. We show the impact of them on disparities in healthy life expectancy between rich and poor countries.

Methods

We use data on individuals' health and attitudes towards health for 67 countries, taken from the World Health Organization's (WHO) World Health Surveys (WHS). To compute healthy life expectancy we combined these with WHO life tables. Like the WHO we use vignette ratings in hierarchical ordered probit (HOPIT) regression models to correct for differences in the use of response scales. The HOPIT yields cross-country comparable estimates of health domain functioning. We modify this approach of the WHO to the measurement of healthy life expectancy in two important respects. First, we do not truncate the resulting health scores to those of the best and the worst vignette in each domain. Second, we retain more inter-individual health variation by combining observed individual health responses with HOPIT outcomes.

Results

We find that truncation of domain scores leads to underestimation of health disparities. In particular, the difference between average healthy life expectancy in the richest and the poorest groups goes from 23.4 to 25.8 healthy life years, when untruncated scores are adopted. Secondly, we show that combining the HOPIT outcomes with original responses to health questions increases health variation within countries, but leads to little changes in countries' average healthy life expectancy.

Conclusions

We introduced modifications to an existing methodology for the measurement of population health. Our results showed that estimates of global variation in population health are sensitive to choices in the methodology applied. Our modifications can be helpful in other studies also applying HOPIT models.

5.1. INTRODUCTION

Improving population health is the defining goal of health systems and as such health attainment is a crucial indicator of the effectiveness of national health policies. Health systems have seen a tremendous improvement in this respect in the past century. Not only has life expectancy increased world-wide, for example American women improved their life expectancy from 71.1 in 1950 to 79.5 in 2000 and are projected to live 87.0 years by 2050. These numbers are even more pronounced for Japanese women, being 60.9, 84.6 and 92.1 years for 1950, 2000 and 2050 respectively (Bongaarts, 2006). Additionally, disease free life expectancy is higher as result of a later onset of morbidities (Fries, 2000). These two developments do not however occur everywhere at the same time nor with the same magnitude (Kalache et al., 2002). Diabetes prevalence for example is higher in developed countries, such as the United States, than in less developed countries such as China. Trends, however, show that these differences are changing. Prevalence in the U.S was 7.4% in 1995 and 2.0% in China, whereas in 2025 these numbers are projected to have grown by 20% and 70% to 8.9% and 3.4% respectively (King et al., 1998). This illustrates how countries vary in the way they progress through the stages of the epidemiological transition (Lopez and Murray, 1998). By comparing levels of health over time and between settings this transition can be monitored and the relative performance of countries' health systems can be assessed.

Comparative studies have traditionally focused on child mortality (Black et al., 2003) or disability free life expectancy (DFLE) (Minicuci et al., 2004). However, more and more studies compare levels of population health using health adjusted life expectancy (HALE), which either uses: i) disability-adjusted life expectancy (DALE) (e.g. Murray and Lopez, 1997; Anderson and Hussey, 2001) based on disability adjusted life years from burden of disease studies or ii) healthy life expectancy -sometimes also referred to as health expectancy - which accounts for cross-population comparability of estimates. Like DFLE, the two HALE alternatives combine estimates of mortality with health related quality of life, but go further by weighting life years based on quality of life estimates (Mathers et al., 2003c).

An important publication using the DALE is the World Health Report (WHR) (2000), in which the World Health Organization (WHO) set out to measure the performance of national health systems, including (population) health attainment. This was the first study to estimate DALE at the country level by combining national mortality data with disability weights from the Global Burden of disease studies. From that year on, the WHO has reported levels of health in its member countries regularly. Importantly, the WHO has also substantially contributed to the further development of the methodology on health expectancy measurement (Murray and Evans, 2003a). This methodology was especially designed to address the concern of cross-country comparability of data, by means of anchoring vignettes and hierarchical ordered probit (HOPIT) regressions (Tandon et al., 2003). Vignettes describe objective levels of health which individuals are

asked to rate in the same scale as they rate their own health. Vignettes are used as anchors for individual interpretation of response scales and combined with own health ratings in the HOPIT analysis.

The investments of the WHO in the availability of country level data and development of estimation techniques have improved the reliability of estimates of health attainment considerably since the WHR 2000. However estimation of healthy life expectancy is still not straightforward and is amenable to improvements. First of all, data on disability levels is still not available for all countries and the most recent data, i.e. the WHS, has to the best of our knowledge only been used in the statistical database of the WHO (World Health Organization, 2012a), but the exact methodology used to arrive at these outcomes has not been published. Secondly, the calculation of healthy life expectancy consists of a number of consecutive steps that have been discussed separately in the literature. Not all steps have been discussed in a single empirical paper, which has limited the transparency of these methods. The efforts of the WHO include the collection of better data, which is reflected in the Multi-Country Survey Study on Health and Responsiveness 2000-2001 (MCSS) (Üstün et al., 2003c) and its follow-up the World Health Surveys 2002-2003 (WHS) (Üstün et al., 2003a).

In this paper, we use the WHS and apply the steps proposed by the WHO for the estimation of healthy life expectancy. We also apply some changes to this WHO approach that either constitute modifications or adaptations. These modifications may help to improve the original methodology, whereas the adaptations are necessary to apply the WHO approach to the WHS. The modifications relate to the way in which the HOPIT model is used to obtain health scores by domain. Although the HOPIT represented a major step forward towards improving cross-country comparability, the way in which the results of this model are used in subsequent steps in the computation of healthy life expectancy requires further attention. First, we propose an alternative method for obtaining health domain scores, which, unlike the standard WHO one, does not assume that all variation in true health is captured by the limited set of observed variables included in the HOPIT. Rather, it assumes that all such unexplained variation in health represents true health differences. Secondly, the WHO rescales the HOPIT outcomes to a zero to one scale to compare and aggregate health domains. This may result in an unnecessary loss of variation, because a large proportion of those estimates are set/truncated to the bounds zero and one. These bounds are determined by the best and worst vignettes in a given health domain which may be far from the best and worst imaginable levels of health in that domain. We propose avoiding such truncation altogether.

In the next section, we describe the WHS and additional data that we use to estimate healthy life expectancy. The third section describes the methods in detail: (i) a stepwise description of the methodology used by the WHO so far, (ii) some adaptations to this methodology that are necessary given the available data, and (iii) our proposed modi-

fications. Section four presents results step by step with and without our modifications. Finally, the fifth section concludes.

5.2. DATA

(a) The World Health Surveys

We use the World Health Surveys (WHS) (World Health Organization, 2002; Üstün et al., 2003a)., a cross-country dataset developed by the WHO. These surveys contain information on household level and in-depth information for one randomly selected household member. The individual questionnaire was used to assess the level of health across countries (World Health Organization, 2002). To create a nationally representative sample, the WHO applied multi-stage stratified cluster sampling of the 18+ population.

(i) Country groupings

We present some results for the following six groups of countries defined by GDP per capita.

0-1000 $: Bangladesh, Congo, Comoros, Ethiopia, Kenya, Lao, Mali, Myanmar, Malawi, Nepal, Chad, Zimbabwe and, Lao

1000-2500 $: Cote d'Ivoire, Ghana, India, Sri Lanka, Mauritania, Pakistan, Philippines, Senegal, Vietnam and South Africa

2500-10000 $: Bosnia Herzegovina, China, Ecuador, Georgia, Guatemala, Morocco, Paraguay, Swaziland, Tunisia, Ukraine Burkina Faso, Dominican Republic, Kazakhstan, Mexico, Mauritius, Malaysia and Russia

10000-25000 $: Brazil, Estonia, Croatia, Hungary, Latvia, Portugal, Slovakia, Slovenia, Turkey, and Uruguay

25000+ $: United Arab Emirates, Austria, Belgium, Czech, Germany, Denmark, Spain, Finland, France, United Kingdom, Greece, Ireland, Israel, Italy, Luxembourg, The Netherlands, Norway and Sweden

(ii) Ratings of individual own health and of hypothetical case vignettes

The WHS contained individual health questions on eight different health domains: mobility, cognition, pain, self-care, interpersonal activities, affect, vision and sleep/energy. Questions were formulated as "Overall in the last 30 days (…) how much difficulty did you have with (…)?" with response categories: "1. None"; "2. Mild"; "3. Moderate"; "4. Severe"; "5. Extreme". Respondents were also asked to rate their overall health (i.e. self-assessed health) through the question: "In general, how would you rate your health today? ",with response categories "1. Very good "; "2. Good"; "3. Moderate"; "4. Bad"; "5. Very Bad".

Furthermore, the survey tried to establish a frame of reference for each respondent with respect to each of the health domains. This was done by means of health vignettes. Vignettes are descriptions of hypothetical situations regarding functioning on the same eight health domains. Individuals are asked to rate these vignettes on the same scale as they use when rating their own health in that domain. An example of a vignette in the domain of mobility is: "[Mary] has no problems with walking, running or using her hands, arms and legs. She jogs 4 kilometers twice a week". The full set of health vignettes used in the WHS can be found in the appendix. The vignette questions were formulated similar to the self-assessment question (e.g. "how much difficulty did [Mary] have with…") but without the reference to the last 30 days. The respondents were asked to imagine the vignette subject to be of the same age and background as him or herself.

(iii) The WHS sample

Later on we link the responses to the health questions in section 5.2(a)(ii) to a set of health determinants: country, age, gender and educational attainment. Education attainment was used to measure socio-economic differences in health and was measured as the number of years a respondent went to school, including higher education. We therefore excluded respondents with missing data on any of the health questions, age, gender or number of education years. Table 5.1 summarizes these characteristics by income group. The survey contained more women (55.6%) than men and the respondents were on average middle-aged (41.2). The number of respondents varied substantially across countries, mainly due to the fact that the samples were larger for the countries where a longer version of the questionnaire was implemented, as opposed to the short version (Üstün et al., 2003a). The average number of years of formal education is higher in the higher income

TABLE 5.1: Descriptive statistics by income groups measured in US $ PPP

	Females				Males			
	N	Age	Years of education	% Bad SRH	N	Age	Years of education	% Bad SRH
Income group 1 ($0 - $1,000)	31,106	38.2	3.9	37%	25,448	39.3	5.3	31%
Income group 2 ($1,000 - $2,500)	27,471	39.0	5.5	39%	25,571	39.1	6.8	32%
Income group 3 ($2,500 - $10,000)	60,225	41.7	7.4	45%	45,043	41.9	7.9	35%
Income group 4 ($10,000 - $25,000)	15,537	43.9	7.7	50%	11,353	44.0	8.5	37%
Income group 5 (> $25,000)	13,705	49.0	11.0	33%	10,262	48.0	8.9	26%
Total	148,044	43.4	8.1	41%	117,677	43.1	8.9	33%

groups and among men. The respondents in the lowest income group are youngest and report health levels that are comparable to the other regions.

(b) Other data sources

The WHO has summarized mortality rates for all its member states by age and gender (World Health Organization, 2007). We used the WHO life tables for 2006 as the inputs for the mortality component of healthy life expectancy. Additionally, we will need an external source on average health utility scores for a national representative sample. For this purpose, we use the data collected by the Measurement and Valuation of Health study (MVH Group, 1995) among the non-institutionalized adult population of the United Kingdom. Respondents reported their functioning on five health domains: mobility, pain, usual activities, self-care and anxiety. These responses were summarized in a health-related quality of life score, which described the distribution of health in the United Kingdom (Dolan, 1997). This distribution was used in our valuation of the different health domain scores of the WHS sample.

5.3. METHODS

The WHO has used healthy life expectancy at birth as a measure of population health. Currently the WHO combines outcomes of the global burden of disease study with those of healthy life expectancy based on household survey data (Mathers et al., 2003a; World Health Organization, 2012a). In this study we focus on the latter. Healthy life expectancy accounts for variation not only in life expectancy at birth (LE), but also in the morbidity experienced throughout life (Mathers et al., 2003c). The household survey based WHO methodology involved four broad steps, which we explain in detail in the remainder of this section. We have to note here that although we refer to this as 'the WHO methodology', there is no single document discussing all of these steps in their entirety. We based our description of this methodology on published papers of healthy life expectancy using the HOPIT methodology (Mathers et al., 2001a, 2003b, 2004) and several chapters on measuring health attainment in Murray and Evans (Mathers et al., 2003a, 2003c; Salomon et al., 2003b, 2003c; Tandon et al., 2003).

(a) – Step 1 – Modeling self-reported health in separate domains, with vignette-correction for reporting heterogeneity – HOPIT models

Self-reported health measures are widely used. In particular, general self-assessed health has been shown to contain valuable information on individual health, in that it is a powerful predictor of mortality, even conditioning on more objective health measures (Idler and Kasl, 1995; Idler and Benyamini, 1997), and of medical care use (Doorslaer and Koolman, 2004). There are however concerns with heterogeneity in the reporting of

health. Individuals in the same, unobserved, health state may rate themselves in different categories on the discrete scale, due to different expectations for their own health, comprehension of survey questions and general cultural factors that may create a tendency towards higher (lower) or extreme (middle) categories. More detailed descriptions of this issue can be found elsewhere (Kerkhofs and Lindeboom, 1995; Groot, 2000; Sadana et al., 2002; King et al., 2003; Lindeboom and Doorslaer, 2004; Etilé and Milcent, 2006).

Aiming at obtaining cross-country comparable health measures, the WHO uses anchoring vignettes, which reveal differences in the use of response scales across individuals. A vignette is essentially a clear cut characterization of the level of ability on a specific domain. The level of ability is fixed in the vignette description, thus the variation in ordinal responses to the vignette questions can be attributed to the variation in the use of cut-points (i.e., to response heterogeneity) (Salomon et al., 2003c; Tandon et al., 2003). The HOPIT combines vignette ratings with self-assessed health to estimate cut-points that vary according to individual characteristics. As a result it is possible to disentangle the reporting effects from true health effects of those characteristics. This generalizes the standard ordered probit model that in turn assumes constant cut-points and therefore is only able to estimate a mixture of true health effects and reporting effects.

More formally, in the first component of the HOPIT model the perceived latent health level of vignette j in domain d, V^{*}_{jdi}, is specified to depend solely on the corresponding vignette being rated and a random, normally distributed error:

$$\text{eq 5.1} \qquad V^{*}_{jdi} = \nu_{jd} + \upsilon_{jdi}, \quad \upsilon_{jdi} \sim N(0, \sigma^2_{vd})$$

This equation contains no individual characteristics as a consequence of the assumption that, up to random measurement error, all respondents understand the vignette description as corresponding to the same level of functioning on a uni-dimensional scale – vignette equivalence. The observed categorical vignette rating, V_{jdi}, relates to V^{*}_{jdi} in the following way:

$$\text{eq 5.2} \qquad V_{jdi} = k \quad \text{if} \quad \tau^{k-1}_{di} \le V^{*}_{jdi} \le \tau^{k}_{di}, \qquad k = 1, \ldots, K,$$

$-\infty < \tau^{1}_{di} < \ldots < \tau^{K-1}_{di} < \infty$. The reporting cut-points are defined as functions of individual characteristics, X_i:

$$\text{eq 5.3} \qquad \tau^{k}_{di} = X_i \gamma^{k}_{d}$$

The second component of the HOPIT defines the latent level of individual own health, H^{*}_{di}, and the observation mechanism that related this latent variable to the observed outcomes, H_{di}:

eq 5.4 $H^*_{di} = X_i\beta_d + \varepsilon_{di}, \quad \varepsilon_{di} \mid X_i \sim N(0, \sigma^2_d)$

eq 5.5 $H_{di} = k \quad \text{if} \quad \tau^{k-1}_{\text{di}} \le H^*_{di} \le \tau^k_{\text{di}}, k = 1, \ldots, K$

The cut-points τ^k_{di} are defined as in eq 5.2, following from the response consistency assumption that any systematic biases in the reporting of own health correspond to those observed in the reporting of the vignettes. It is assumed that the error terms in the vignette. In this model, only σ_d/σ_{vd} is identified but not σ_d and σ_{vd} separately. For this reason, we normalize one of them ($\sigma^2_{vd} = 1$).

We considered as individual characteristics in the vector X_i indicators of country and gender, a second order polynomial for age, the number of education years and the interaction between education years and country. These variables enter the HOPIT model both in the own health equation (eq 5.4) and in the cut-point equations (eq 5.3). This makes it possible to separately identify health and reporting effects of those variables, aiming at obtaining adjusted health scores purged of reporting heterogeneity to be used in the next step.

(b) – Step 2 – Obtaining vignette adjusted health scores by domain from HOPIT models

For each health domain from the HOPIT procedure the WHO estimates a linear index for each respondent using estimated eq 5.4:

eq 5.6 $\hat{H}^*_{di} = X_i\hat{\beta}_d$

The WHO transforms the latent linear indices for each domain so as to range from the best to the worst possible domain performance. First, anchors are selected to determine what constitutes best and worst possible domain performance on the latent scale. Several anchors have been used in HOPIT studies so far. We assume here that the WHO used the predicted latent health levels of the best and worst vignettes, as described in the paper which forms the basis for the WHO approach (Salomon et al., 2003a) in the next step.

(c) – Step 3 – Aggregation of health domain scores into a general health utility score

The WHO has investigated the potential for using multi-attribute utility function to combine the performance on different domains into a single health score. For this purpose Salomon et al. (2003b) have provided a set of valuation functions. Since none of the papers on healthy life expectancy contain a detailed discussion of this step, we assume the WHO still relies on valuation functions such as those of presented by Salomon et al. These valuation functions aggregate the truncated latent health scores of all domains, obtained

in step 2, after they have been further rescaled into a zero to one interval, through a linear transformation: $\hat{H}[0,1]_{di}$.[61] They are defined as follows:

eq 5.7 $$\hat{HU}_i = \hat{\alpha}_0 + \sum_{d=1}^{D} \hat{\alpha}_d \hat{H}[0,1]_{di}$$

where $\hat{\alpha}_d$ is the valuation weight for domain d and $\hat{\alpha}_0$ is a constant, which have been estimated by Salomon et al. (2003b). These have been obtained by regressing visual analogue scale (VAS) scores of fixed health states on estimated HOPIT scores for the same health states. A VAS score is a valuation between 0 and 100 of a health state directly provided by a respondent. These models have been estimated for MCSS data which did not cover the same health domains as the WHS. What is more, the set of valuations that was collected in the WHS is not suitable to be analyzed in a similar fashion. The WHO intended to avoid direct valuations of health states and chose not to collect any direct valuations of health states. Therefore, we need to derive valuations of the HOPIT based domain scores in another way. The WHO valuation function aggregates domain scores directly into health utilities. Our adapted procedure does this in two sub-steps.

1) Adaptation 1: Different aggregation function for domain scores

We use the self-assessed overall health of respondents on a five point scale to elicit valuations of different health domains. As the respondents also answered the eight health domain question, this is an implicit valuation of their combined performance on all domains. We uncover these implicit valuations by using a standard ordered probit regression of the overall health on the health domains. This approach assumes that all variation in health is reflected in the ordinal responses to the health question. The self-rated domain questions entered the regression through dummy variables for each different category to allow for the differential impacts on health across the domain scale. So, the latent overall self-assessed health level is specified as follows:

eq 5.8 $$SAH_i^* = \delta_0 + \sum_{d=1}^{8} \sum_{k=2}^{5} \delta_k^d h_{ki}^d + u_i, \quad u_i \sim N(0,1)$$

where h_{ki}^d are indicators of whether $H_{di}=k$. Overall self-assessed health relates to the latent index as:

$$SAH_i = k \text{ if } \lambda^{k-1} \leq H_i^* \leq \lambda^k, k = 1,\ldots,5, -\infty < \lambda^1 \leq \ldots \leq \lambda^{K-1} < \infty.$$

61. This yields scores on a zero to one scale, but the scale itself is still ordinal. As a result, further interpretation of differences in domain performance between individuals or over time is still not possible. For example, living two years with a mobility domain score of 0.2 does not have to be valued equally to living one year with a mobility domain score of zero and one year with 0.4, even when ignoring discounting.

We used estimates of this model to aggregate predicted latent health scores across domains. In particular, we combined estimated standardized coefficients of this ordered probit regression with the domain scores from step 2 in the following way:

$$\text{eq 5.9} \quad \hat{H}_i^* = \sum_{d=1}^{8} \sum_{k=1}^{5} h_{ki}^d \frac{\hat{\delta}_k^d \hat{\sigma}_k^d}{\hat{\sigma}^h} \tilde{H}_{di}^*$$

where $\hat{\delta}_k^d$ are the original estimated coefficients from the ordered probit analysis; and the standard deviation of the specific domain ($\hat{\sigma}_k^d$) divided by the standard deviation of the dependent variable ($\hat{\sigma}^h$) describes the standardization of the coefficient; and $\tilde{H}_{di}^*$ is the individual domain score of individual *i* on domain *d* from step 2, standardized to have a mean of zero and standard deviation of 1 using sample statistics.

2) Adaptation 2: Transforming latent health scores into health utilities

The valuating function used by the WHO aggregates the health domains, but also transforms the scores to a utility score. This transformation allows the aggregated scores to be combined with mortality data in the next step. Our adapted aggregation function results in a general health score that still has to be transformed into a health utility score. Therefore an additional adaptation is needed. This maps EuroQol data of the general population of the United Kingdom (Kind et al., 1998) to our aggregated health scores. Using the EuroQol "tariff" for the UK (Dolan, 1997) the Kind et al. study provides a distribution of health utility in the UK.

We assume that health utility is distributed similarly in the UK WHS sample. Consequently, we can assume that the health utility of the one percent unhealthiest UK respondents in the WHS, ranked by the scores of eq 5.9, is the same as that of the one percent unhealthiest respondents in the Kind et al. study. Therefore, the one percent unhealthiest individuals in the UK WHS sample are assigned that health utility score. So for example if one percent of the UK WHS sample has an aggregated health score of 2.3 or more and one percent of the Kind et al. study has a health utility of 0.43 or lower, all WHS respondents with an aggregated health score of 2.3 or higher are assigned a health utility of 0.43. For countries other than the UK all respondents with a score of 2.3 or more are also assigned a health utility of 0.43, but this does not necessarily equal one percentage of the respondents in those countries. This mapping is repeated for all other centiles of the UK WHS sample and EuroQol sample. This procedure results in the individual health utility scores $\hat{HU}_i$.

We should note that the choice to use a country as a reference country is not entirely innocuous. First, the whole WHS sample will be assigned a health utility based on UK valuations of health states. If another country's health utilities would be used, different absolute health utilities would be found in each country, but the same relative differences would exist. Secondly, in the end only one hundred different health utilities are applied

in the whole WHS sample. By design each health utility is applied an even number of times in the UK WHS sample. Consequently, a ceteris paribus improvement in responses to a health domain question will not yield a higher average health utility in the UK. However, average health utility for UK men and UK women are not fixed. If for example UK women improve their health, their health scores will be mapped to higher health utilities and consequently men's to lower health utilities, because still one percent of the respondent is assigned the one percent lowest health utility. Changes in health in other countries will be reflected in the average health utility in that country, because it need not be that each health utility is assigned to one percent of the respondents in that country. As we will see later, it is important to keep this in mind when we interpret the impact of our modifications in section 5.4.

(d) – Step 4 – Estimation of healthy life expectancy combining health utility scores with mortality data at different ages

Once health utilities are derived the resulting individual health utilities are combined with the mortality data to obtain estimates of healthy life expectancy. The WHO does this using Sullivan's method (1988) which first takes averages of health utilities at different ages for different sexes and then multiplies these with the number of years lived between those ages (Mathers et al., 2001b). The age ranges used depend on the life table data. Generally, five year wide averages are used except in the youngest ages. The WHS did not cover individuals under 18, but for this study it is necessary to also estimate health levels during childhood. Therefore, we have estimated health levels for children using the scores of respondents aged 18 and 19. Since the dataset included very few respondents older than 75 a larger age-interval grouping was also used to determine the average health levels for this group of respondents. In this grouping we used the health levels of all respondents above the age of 75. Average level of health in country *c* for gender *g* within age range *x* is then defined as:

$$\text{eq 5.10} \quad HU_x^{cg} = \frac{\sum_{i=1}^{N_x^{cg}} HU_i}{N_x^{cg}}$$

Subsequently healthy life expectancy estimates are calculated by combining survival data with estimates of average levels of health as follows:

$$\text{eq 5.11} \quad HLE^{cg} = \sum_{x=0}^{22} L_x^{cg} * HU_x^{cg}$$

where L_x^{cs} is the expected number of years lived during each age-interval x in country c for gender s.

5.4. TWO POSSIBLE MODIFICATIONS TO THE WHO METHODOLOGY

In *Step 2* of the WHO methodology health scores are predicted using eq 5.6. In this section we propose two modifications to the procedure for obtaining health scores per domain. The first one relates to the use of eq 5.6 and the second one to the subsequent truncation of these scores.

(a) Modification 1: Alternative prediction of latent health domain scores from HOPIT model

A first drawback of the WHO method is that all individuals with identical personal characteristics are assigned an equal score, even when they report different answers to the domain question. This reduces variation in individual predictions of domain functioning. The linear index will generally not capture all variation in health in the respective domain. It is likely that some unexplained variation is due to insufficient control variables in our model. Examples of possibly relevant variables excluded from the analysis could be income, working- or living conditions, availability of health care and individual frailty. In our model, especially the country dummies and the country specific effects of education should capture large shares of the variation associated with such variables. For example, income will impact health levels of individuals, but this will in part be captured by education. Other unobserved factors may have a larger proportion of their variation not captured by our observed covariates.

Our approach uses the estimated HOPIT model in a different way, which does not rely solely on that predicted linear health index. We use information on observed categorical outcomes. In particular, we obtain predicted values in the latent health scale as "midpoints" in the interval that corresponds to those observed outcomes. The HOPIT model is used to estimate these intervals, which are individual specific, and determined by the cut-points defined in eq 5.3. This means that individuals with the same observed outcome but with different values of X_i will have different predicted latent health scores, reflecting reporting heterogeneity. In other words, in that case, the same observed outcome actually corresponds to a different interval in the latent health scale.

We maintain the normality assumption of eq 5.4. For a given individual with observed health outcome in domain d, H_{di}, equal to k, the predicted latent health score in that domain is the median score between cut-points $\hat{\tau}_{\mathrm{di}}^{k-1}$ and $\hat{\tau}_{\mathrm{di}}^{k}$ of a normal distribution with mean $\hat{\mu}_{di} = X_i\hat{\beta}_d$ and standard deviation of $\hat{\sigma}_d$, all estimated by the HOPIT model. These estimated median scores can then be obtained as:

$$\text{eq 5.12} \quad \hat{H}^{*}_{di} = \hat{\mu}_{di} + \hat{\sigma}_d \times \Phi^{-1}\left\{\frac{1}{2}\left[\Phi\left(\frac{\hat{\tau}^{k-1}_{di} - \hat{\mu}_{di}}{\hat{\sigma}_d}\right) + \Phi\left(\frac{\hat{\tau}^{k}_{di} - \hat{\mu}_{di}}{\hat{\sigma}_d}\right)\right]\right\}$$

Our approach retains more of the variation in observed health outcomes than the standard WHO approach relying on the predicted linear index. This is made under the

assumption that all variation in observed health outcomes that is left unexplained by eq 5.4 can be attributed to variation in true health, rather than to unobserved reporting heterogeneity. Using the predicted linear index, on the other hand, assumes the opposite, and attributes all unexplained variation to unobserved reporting heterogeneity, relying only on the health variation that can be explained by observed covariates.

(b) Modification 2: No truncation of health levels that are better/worse than the best/worst vignette

The disadvantage of the rescaling approach in step two of the WHO methodology is that any predicted scores below/above the chosen anchors are truncated to the respective scores. By truncating the scale in this way some of the variation in predicted HOPIT scores will be lost. We propose to not to truncate the outcomes of the HOPIT analysis. In this way all variation remains when determining domain and ultimately overall health performance.

Truncation of the HOPIT scale will hide differences in true health, especially in health domains where many respondents report better (worse) functioning than the best (worst) vignette. If for example in both Spain and Germany many individuals are predicted to see better than the vision described in the best vision vignette, for all these respondents the vision score will be rescaled to 0 (i.e. least problems). However, if the respondents from Germany on average were further away from the level of vision described in the vignette than those in Spain, then this difference in vision between the two countries disappears with the truncation, as both are assigned the same score. This problem will probably occur most often around the best vignette and less often around the worst vignette, since there are relatively few individuals performing that poorly in a general population. In addition, the problem will not be equally important for each domain. The importance will be determined both by the health described in the vignette and number of respondents encountering problems.

5.5. RESULTS

In this section, we present results for healthy life expectancy for all countries considered in the analysis, following the WHO methodology for steps 1, 2 and 4. As explained above: i) in step 3, we were not able to apply the WHO methodology and so applied only our necessary adaptations of it; ii) in step 2, we propose some modifications to the WHO methodology and present/compare here results of both the original and the modified approaches from step 2 onwards.

(a) – Step 1 – Vignette correction for reporting heterogeneity – HOPIT models

We observe large country differences in the assessments of identical health levels represented by the vignettes. Take the example, of the pain vignette "[Mark] has pain in his knees, elbows, wrists and fingers, and the pain is present almost all the time. It gets worse during the first half of the day. Although medication helps, he feels uncomfortable when moving around, holding and lifting things." 80% of Nepalese women between 40 and 45 consider that this vignette corresponds to severe or extreme pain, while this is only the case of the 65% of their Philippine counterparts. Large differences are also found in the ratings of this and other vignettes across countries, and within countries across socio-demographic groups.

It is this variation in vignette ratings that the HOPIT model exploits to identify reporting heterogeneity and in turn to purge this reporting heterogeneity from individual self-reported levels of functioning in each health domain. To save on space, we do not present here full results of HOPIT estimation. Wald tests of significance of all and of groups of variables in the reporting cut-points (eq 5.3) show the importance of taking reporting heterogeneity into account. Each of these Wald tests showed significant reporting differences (all p-values smaller than 0.001).

(b) – Step 2 – Domain functioning

We first present results of the WHO approach for obtaining vignette adjusted latent health domain scores from the estimated linear health index of the respective HOPIT model. In the WHO methodology this involved truncating the estimated individual linear indices using those estimated for the best and worst vignettes in each domain. Table 5.2 shows that truncating to the best vignette is much more common than to the worst vignette. Additionally, the number of truncated observations differs considerably by domain. For example, six times as many scores are truncated in the personal relations domain than in the affect domain. In the personal relationships domain, almost all scores are recoded. This depends largely on the descriptions in the used vignettes[62]: the better the health described in the vignette, the fewer respondents will be in better health. Our *modifica-*

62. The best vignette used for the personal relations domain was: Charlie can join in any community activities that interest him, whenever he wants to, without any restrictions. He gets on well with everybody and enjoys meeting new people". The worst vignette on the personal relations domain was: "Amala suffered a stroke three months ago. Her friends do not come and visit anymore as Amala cannot communicate with them. She is constantly upset and shouts at her family members which causes them to avoid her."

The best vignette used for the affect domain was: "Ken loves life and is happy all the time. He never worries or gets upset about anything and deals with things as they come." The worst domain on the affect domain was: "Vivian has already had five admissions into the hospital because she has attempted suicide twice in the past year and has harmed herself on three other occasions. She is very distressed every day for the most part of the day, and sees no hope of things ever getting better. She is thinking of trying to end her life again."

TABLE 5.2: Number of observations recoded to best (0) and worst (1) health in WHO approach

	Number of observations rescaled		Percentage rescaled to best vignette by GDP in 2010 US$ PPP				
	To best	To Worst	0 - 1000	1000 - 2500	2500 - 10000	10000 - 25000	>25000
Mobility	142,960	1	51%	30%	65%	54%	63%
Self-care	244,621	0	91%	86%	95%	93%	95%
Pain	77,240	0	39%	22%	29%	19%	35%
Cognition	94,132	0	41%	22%	38%	31%	47%
Personal relations	259,076	0	96%	95%	98%	100%	100%
Seeing	213,368	94	82%	76%	80%	80%	88%
Sleeping	107,966	0	44%	39%	48%	21%	28%
Affect	42,822	0	19%	17%	13%	7%	30%

tion 2 simply consists of using untruncated estimated individual linear indices. Figure 5.1 compares results of the WHO Step 2 with those obtained by introducing modification 2, for two health domains (personal relationships and affect) by gender and region (divided in poor and rich countries) and across age groups (x-axis). Here, we have selected for illustrative purposes the domains for which truncating had the greatest (personal relationships) and smallest impact (affect).

For personal relationships we find that before truncation (squares) there are differences by gender, region, at younger ages, which largely disappear after truncation. This is because *all* individuals at young ages, no matter if they are young women from poor countries (black x's) or males from rich countries (gray x's) are predicted to be in better health than the best vignette. Horizontal lines representing the best and worst vignettes

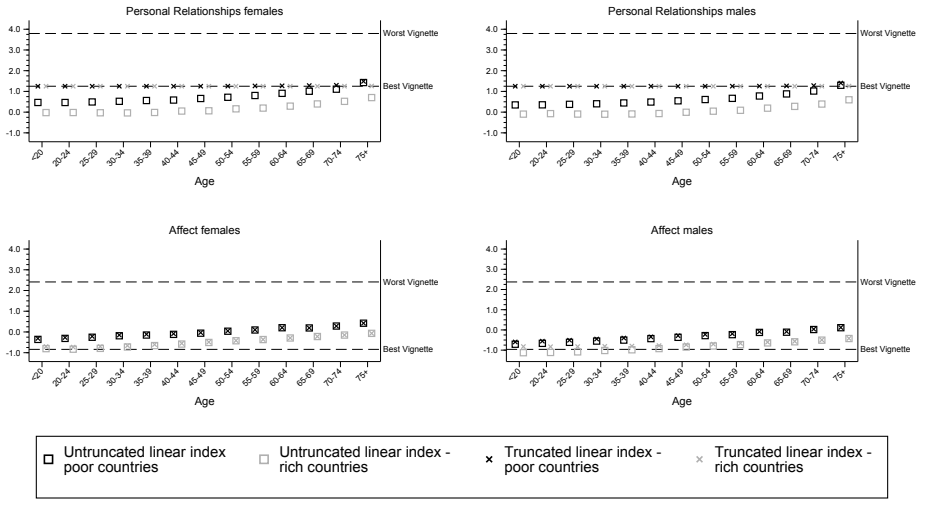


FIGURE 5.1: Effects of truncating for the domains of personal relations and affect

have been added to each panel. At older ages, reflecting worsening of functioning, the impact of truncation becomes smaller. This is because some individuals in those groups are expected to have predicted scores of problems with personal relationships worse than the best vignette, which in turn will not be truncated (while at the same time predicted scores are still generally better than those of the worst vignette). For affect, the truncated results still show differences in domain functioning, between age, gender and income groups. There is however still less variation in the truncated than in the untruncated distribution, especially at younger ages, reflecting the fact that some individuals are predicted to have better functioning in this domain than the best vignette. These young individuals then have their score truncated which leads to a worse average score in their group, and closer to subsequent older groups, than that obtained without truncation.

We now look at the impact of our *modification 1* in the computation of latent health scores by domain. This consists of using median scores of individual intervals based on observed responses (using eq 5.12rather than simply using the estimated linear health index of the respective HOPIT (eq 5.6). In Figure 5.2 we provide kernel density using both methods for the same two domains as before (without truncation). The linear index (black line) shows a rather smooth distribution of performance on the domain of affect and somewhat less for personal relations. This difference is a result of the effects and distribution of the determinants in the HOPIT analysis. If we turn to the gray lines, i.e. the scores based on individual responses, we find a larger spread of the scores, which is not surprising given that those with a good (bad) linear index and picking the best (worst) health response

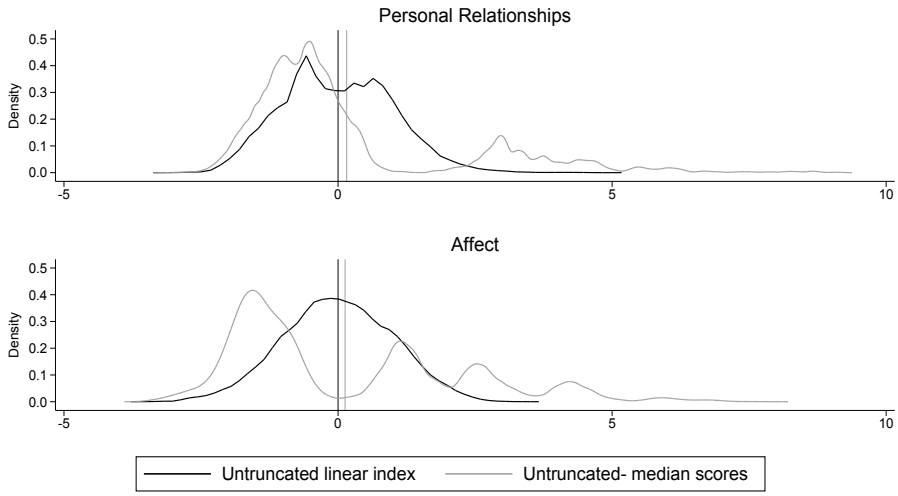


FIGURE 5.2: Kernel density of HOPIT based domain scores using the linear index and median based health scores for the domains of personal relations and affect

will have an even better (worse) score using the individual cut-points. The distribution of individual cut-point based scores clearly shows some additional spikes indicating the different possible responses. Since scores are centered around these spikes, it is plausible that the individual responses are more important for a domain score than an individual's background characteristics. If the latter were the case the responses would be centered more around the mean. Finally the vertical bars show that the average domain scores using the individual cut-points are slightly higher (i.e. worse functioning) than those based on the linear index.

(c) – Step 3 – Aggregated health levels

To show the importance of the selected reference country, Table 5.3 shows health utilities for 10 countries including the UK. What immediately becomes clear from the table is that no matter which method is used, the average health utility in the UK remains the same. The intuition behind this finding has been given in section 5.3(a)(iii)2). For the other countries the average health utilities do change, although for some more than for others. Since the UK utilities remain constant, these changes need to be interpreted as relative to the UK. We will nevertheless still see changes in the healthy life expectancy scores of the UK, because the distribution of the health utilities over males and females and different age groups does differ between the methods. This is important because of the way in which health utilities by age-gender group and life table data are combined in Sullivan's method.

TABLE 5.3: Average health utility for the United Kingdom and 9 other countries with and without proposed modifications

	Untruncated Linear Index	Linear Index with Truncating (WHO)	Median scores
United Kingdom	0.895	0.895	0.895
Brazil	0.846	0.820	0.846
Comoros	0.772	0.904	0.778
Hungary	0.922	0.938	0.911
Ireland	0.975	0.985	0.959
Kazakhstan	0.902	0.952	0.886
Netherlands	0.896	0.862	0.914
Senegal	0.717	0.805	0.631
Vietnam	0.966	0.992	0.936
South-Africa	0.747	0.849	0.714

(d) – Step 4 – Final results of healthy life expectancy

Figure 5.3 shows how healthy life expectancy differs, on average, between the regions. The bars show a clear gradient in health attainment across income groups. A second clear pattern is that the impact of the different methods is much larger for lower income

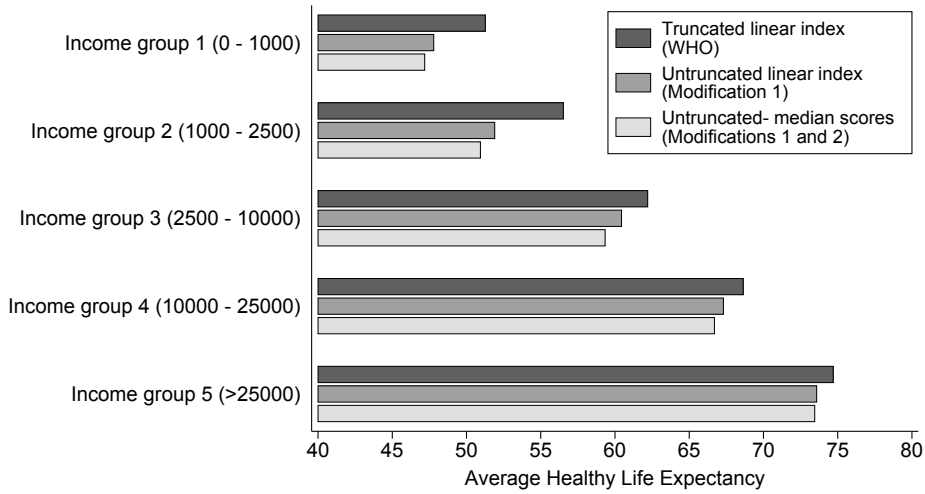


FIGURE 5.3: Estimates of healthy life expectancy by GDP country with and without modifications

than for higher income countries. This is because the United Kingdom is in this group and the health levels of other countries in these groups are similar to those of the United Kingdom.

The most important thing to note from Figure 5.3 is the effect of truncating domain health scores on measured healthy life expectancy, *modification 2*. Table 5.2 above showed that truncation is more common in the rich countries. For example, in the affect domain the share of recoded observations is approximately 50% larger in the group of rich country than in the poorest countries and truncation is most of the times to the score of the best vignette. The extent to which a higher number of truncated observations results in a relatively worse estimate of healthy life expectancy depends also on how far apart the original own health score was from that of the best vignette. This translates into differences in healthy life expectancy as shown in Figure 5.3. The difference between the richest and poorest group of countries increases from 23.4 to 25.8 healthy life years, a 10.2% increase. The direction of this effect is what one would expect. Truncating improves health of the poorer countries, relative to the UK, because there are fewer observations truncated to the best vignette here than in the rich countries. Since the absolute levels of healthy life expectancy in the rich countries are more or less fixed - if they do not differ too much from the UK – these relative deteriorations will be reflected in higher healthy life expectancies in the poor countries.

Our proposed *modification 1* has a much smaller impact on measured healthy life expectancy. The third bar in Figure 5.3 is based on the results using untruncated median scores and shows small differences compared to the second bar, even for poorer countries. The difference between the richest and poorest group of countries does increases further from

25.8 to 26.3 healthy life years, a 1.9% increase. We did not necessarily expect a different result regarding country averages. Figure 5.2 showed that the average domains scores were similar in both methods. Consequently, overall health scores (from eq 5.9) and healthy life expectancies (eq 5.11) will also be similar. It is in the assessment of health inequality within these groups that the use of median vs. linear index scores has the most potential.

Finally, we show the results of health adjusted life expectancy for each country in Table A. 5.1 for reference. The analyses of the impact of our modifications apply as described above, but the results for individual countries do not always follow the same pattern as that of the rest of the income group. For example, Guatemala's healthy life expectancy improves with more than 3 when scores are not truncated, whereas on average it drops within the third group of countries, to which Guatemala belongs.

5.6. DISCUSSION

We calculated estimates of population health in the 67 countries participating in the World Health Surveys. Our methods add to the current methodology in two ways. First, our results do not truncate outcomes of the hierarchical ordered probit model. Secondly, we used an alternative outcome of the HOPIT analysis to measure individual performance and retain variation in health scores. In order to be able to calculate healthy life expectancy we also had to make some adaptations to the WHO methodology. First, we measured preferences over health states using implicit valuation within our primary data source. Secondly, we used external data to map health scores from a latent scale to a health utility scale.

We showed that truncation of HOPIT scores does not only lead to a loss of variation in health outcomes at the individual level, but it also estimates smaller discrepancies in health attainment between rich and poor countries. In particular, not truncating health domain scores increases the difference between health life expectancy of the richest and poorest groups with 10.2%.The impact of the modification depends heavily on the description of health in the vignettes. This is because having as a best/worst vignette in a given domain a situation which does not correspond to the highest/lowest imaginable level of health in that domain will increase the proportion of observations for which the health score is (unduly) truncated.

Our second modification relies on the assumption that not all health variance is captured in the specification of the HOPIT and therefore uses observed individual responses in the estimation of individual domain performance. In particular, the individual response is translated into the median score within the observed category, but the boundaries of the respective interval vary with individual characteristics, as estimated in the HOPIT model. This modification has a small impact on the outcomes of this paper. Its potential applicability, however, extends beyond the measurement of average health

attainment. It will for example play a more important role in analyses of health inequalities between individuals within countries. Another potentially important application is the use of vignette adjusted health scores as explanatory variables in models of health care use (as in Bago d'Uva et al., 2011) and labor market participation (e.g. Kapteyn et al., 2009). Simply using the predicted linear index requires either exclusion restrictions in the HOPIT model (Bago d'Uva et al., 2011) or restrictions such as independence between the two parts of the joint model (Kapteyn et al., 2009). These are not required in the case of median scores since these are not simply linear functions of the (same) explanatory variables.

We found evidence of large variation in health attainment between countries around the globe, which becomes larger after our modifications. Estimates of health adjusted life expectancy are below 50 in the poorest group of countries and close to 75 in the richest. These differences are driven both by health utility and life expectancy. There are also substantial differences between rich countries (Table A. 5.1). Those inequalities might provide more relevant information for policy than those between the rich and the poor. Information about best practices (e.g. Norway for the developed countries) on prevention, access to health care, organization of care and health system characteristics can help other countries, if comparably wealthy, to improve population health.

We tried to show the effects of both separate modifications as clearly as possible, but were sometimes impaired by the lack of full details that have been published on the WHO methodology. For example, the way in which the HOPIT was exactly specified was not always mentioned, the set of explanatory variables might explain differences in health utility. What is more, the application of the methodology to the WHS has not been discussed at all. We provide these and other details to be transparent and make our results comparable to those of others (Boerma et al., 2010).

The adaptations we have had to make to estimate healthy life expectancy will influence our outcomes. Our aggregation of health domains relies on implicit valuations of health and assumes all relevant variation is captured in the ordered responses to health questions. In contrast to the models estimated by Salomon et al. (2003b) the outcomes of this aggregation procedure do not represent health utilities. To derive utility scores, our analysis relies on external data. The applied mapping procedure has at least one big advantage compared to previous proposals. A mapping procedure like applied here allows for a non-linear transformation of the HOPIT scores. The previous methods, like valuation functions, all made assumptions about constant marginal returns in health utility of a unit increase on the latent health variable.

Methods for the measurement of population health are being developed and need to be further refined in the future. Population health measurement studies such as this one should already be performed regularly to enable monitoring of the effectiveness of health systems, and to evaluate health system reform.

APPENDIX

Mobility

Vignette 1: Rob is able to walk distances of up to 200 metres without any problems but feels tired after walking one kilometre or climbing up more than one flight of stairs. He has no problems with day-to-day physical activities, such as carrying food from the market.

Vignette 2: Vincent has a lot of swelling in his legs due to his health condition. He has to make an effort to walk around his home as his legs feel heavy.

Vignette 3: Mary has no problems with walking, running or using her hands, arms and legs. She jogs 4 kilometres twice a week.

Vignette 4: Anton does not exercise. He cannot climb stairs or do other physical activities because he is obese. He is able to carry the groceries and do some light household work.

Vignette 5: David is paralyzed from the neck down. He is unable to move his arms and legs or to shift body position. He is confined to bed.

Affect

Vignette 1: Jan feels nervous and anxious. He worries and thinks negatively about the future, but feels better in the company of people or when doing something that really interests him. When he is alone he tends to feel useless and empty.

Vignette 2: Vivian has already had five admissions into the hospital because she has attempted suicide twice in the past year and has harmed herself on three other occasions. She is very distressed every day for the most part of the day, and sees no hope of things ever getting better. She is thinking of trying to end her life again.

Vignette 3: Roberta feels depressed most of the time. She weeps frequently and feels hopeless about the future. She feels that she has become a burden on others and that she would be better dead.

Vignette 4: Henriette enjoys her work and social activities and is generally satisfied with her life. She gets depressed every 3 weeks for a day or two and loses interest in what she usually enjoys but is able to carry on with her day to day activities.

Vignette 5: Ken loves life and is happy all the time. He never worries or gets upset about anything and deals with things as they come.

Personal relationships

Vignette 1: Johanna gets on well with the people she knows but has no close friends. She has not spoken to her mother in 5 years and does not want to see her. Because of this tension, her family usually excludes her from family gatherings.

Vignette 2: Nobu is blind and lives in a remote rural area. His family does not allow him to leave the house because they fear he will get hurt. His family tells him that he is a burden to them. Their criticism upsets him and he cries.

Vignette 3: Elizabeth has difficulty climbing up and down the stairs and walking. She is not able to go out as much as she would like to but has many friends who come and visit her at home. Her friends find her a source of great comfort.

Vignette 4: Amala suffered a stroke three months ago. Her friends do not come and visit anymore as Amala cannot communicate with them. She is constantly upset and shouts at her family members which causes them to avoid her.

Vignette 5: Charlie can join in any community activities that interest him, whenever he wants to, without any restrictions. He gets on well with everybody and enjoys meeting new people.

Pain

Vignette 1: Mark has pain in his knees, elbows, wrists and fingers, and the pain is present almost all the time. It gets worse during the first half of the day. Although medication helps, he feels uncomfortable when moving around, holding and lifting things.

Vignette 2: Steve has pain in the neck radiating to the arms that is not relieved by any medicines or other treatment. The pain is sharp at all times and keeps him awake most of the night. During the day the pain has made him completely incapacitated. It has necessitated complete confinement to the bed and often makes him think of ending his life.

Vignette 3: Laura has a headache once a month that is relieved one hour after taking a pill. During the headache she can carry on with her day to day affairs

Vignette 4: Isabelle has pain that radiates down her right arm and wrist during her day at work. This is slightly relieved in the evenings when she is no longer working on her computer.

Vignette 5: Katie cannot remember when she last felt pain as this has not happened for the last several years now. She does not experience any pain even after hard physical labor or exercise.

Vision

Vignette 1: Jennifer only reads if the text is in very large print, such as 10 lines per page. Otherwise she does not read anything. Even when people are close to her, she sees them blurred.

Vignette 2: Norman needs a magnifying glass to read small print and look at details on pictures. He also takes a while to recognize objects if they are too far from him.

Vignette 3: Antonio can read words in newspaper articles (and can recognize faces on a postcard size photograph). He can recognize shapes and colors from across 20 meters but misses out the fine details.

Vignette 4: Sebastian cannot detect any movement close to the eyes or even the presence of a light.

Vignette 5: Hector can read words in newspaper articles (and can recognize faces on a postcard size photograph). He can recognize familiar people's faces all the time and picks out most details in pictures from across 20 meters.

Sleeping

Vignette 1: Noemi falls asleep easily at night, but two nights a week she wakes up in the middle of the night and cannot go back to sleep for the rest of the night. On these days she is exhausted at work and cannot concentrate on her job.

Vignette 2: Mark falls asleep every night within five minutes of going to bed. He sleeps soundly during the whole night and wakes up in the morning feeling well-rested and feels full of energy all day.

Vignette 3: Paolo has no trouble falling asleep at night and does not wake up during the night, but every morning he finds it difficult to wake up. He uses an alarm clock but falls back asleep after the alarm goes off. He is late to work on four out of five days and feels tired in the mornings.

Vignette 4: Damien wakes up almost once every hour during the night. When she wakes up in the night, it takes around 15 minutes for her to go back to sleep. In the morning she does not feel well-rested and feels slow and tired all day

Vignette 5: Daniel takes about two hours every night to fall asleep. He wakes up once or twice a night feeling panicked and takes more than one hour to fall asleep again. Three to four nights a week he wakes up in the middle of the night and cannot go back to sleep for the rest of the night. He is fatigued all day, every day and misses work several times a week. He cannot take part in sports or social activities.

Self care

Vignette 1: Helena pays a lot of attention to the way she looks. She requires no assistance with cleanliness, dressing and eating.

Vignette 2: Anne takes twice as long as others to put on and take off clothes, but needs no help with this. Although it requires an effort, she is able to bathe and groom herself, though less frequently than before. She does not require help with feeding.

Vignette 3: Sandra lives on her own and has no relatives or friends nearby. Because of her arthritis, she is house-bound. She often stays all day in the same clothes that she has slept in, as changing clothes is too painful. A neighbour helps her wash herself.

Vignette 4: Victor requires no assistance with cleanliness, dressing and eating. He occasionally suffers from back pain and when this happens he needs help with bathing and dressing. He always keeps himself tidy.

Vignette 5: Sue is quadriplegic and must be washed, groomed, dressed and fed by somebody else.

Cognition

Vignette 1: Theo cannot concentrate for more than 15 minutes and has difficulty paying attention to what is being said to him. Whenever he starts a task, he never manages to finish it and often forgets what he was doing. He is able to learn the names of people he meets but cannot be trusted to follow directions to a store by himself.

Vignette 2: Peter does not recognize even close relatives and gets lost when he leaves the house unaccompanied. Even when prompted, he shows no recollection of events or recognition of relatives. It is impossible for him to acquire any new knowledge as even simple instructions leave him confused.

Vignette 3: Sue can find her way around the neighborhood and know where her own belongings are kept, but struggles to remember how to get to a place she has only visited once or twice. She is keen to learn new recipes but finds that she often makes mistakes and has to reread several times before she is able to do them properly.

Vignette 4: Malcolm can concentrate while watching TV, reading a magazine or playing a game of cards or chess. He can learn new variations in these games with small effort. Once a week he forgets where his keys or glasses are, but finds them within five minutes.

Vignette 5: Rob is very quick to learn new skills at his work. He can pay attention to the task at hand for long uninterrupted periods of time. He can remember names of people, addresses, phone numbers and such details that go back several years.

TABLE A. 5.1: Healthy life expectancy by country using three methods

	Untruncated linear index (modification 1)			Truncated linear index (WHO)			Untruncated median score (Modifications 1 + 2)		
	females	males	both	females	males	both	females	males	both
Austria	77.0	74.5	75.6	79.1	75.5	77.2	77.2	73.9	75.5
Bangladesh	46.4	45.8	45.8	53.0	49.8	51.3	44.2	48.6	46.3
Belgium	74.7	72.0	73.2	77.0	72.9	74.9	74.5	72.8	73.7
Bosnia and Herzegovina	71.0	68.7	69.9	75.2	71.0	73.2	68.2	64.3	66.3
Brazil	62.7	59.0	60.7	61.9	57.1	59.4	60.9	60.1	60.4
Burkina	45.8	44.7	45.3	47.3	45.2	46.2	44.8	43.6	44.2
Chad	35.0	34.3	34.5	40.1	38.6	39.2	33.2	34.4	33.7
China	69.0	68.8	68.9	72.5	70.8	71.6	68.7	68.1	68.4
Comoros	52.1	50.2	51.2	61.1	57.7	59.4	51.9	50.2	51.0
Congo	41.2	41.0	41.0	47.9	47.8	47.8	39.2	36.7	38.2
CotedIvoire	42.3	42.5	42.6	47.9	45.5	46.7	41.6	41.0	41.5
Croatia	69.8	67.5	68.5	71.6	68.2	69.8	68.8	66.4	67.5
Czech	71.7	68.9	70.3	74.1	70.2	72.1	71.1	68.4	69.8
Denmark	75.2	73.2	74.2	73.9	71.6	72.7	75.7	73.2	74.4
Dominican	62.7	57.8	60.0	62.3	56.1	58.8	58.6	57.3	57.8
Ecuador	63.5	61.0	62.0	66.1	62.2	64.0	62.4	61.6	61.8
Estonia	72.0	64.1	68.0	73.8	64.9	69.4	71.7	64.0	67.9
Ethiopia	49.3	48.9	49.1	54.0	51.8	52.9	50.6	49.0	49.8
Finland	78.3	74.6	76.4	78.1	73.7	75.9	79.1	73.8	76.4
France	73.0	69.7	71.5	73.8	69.5	71.7	70.7	70.6	70.6
Georgia	65.9	62.0	64.0	69.3	63.8	66.5	64.5	60.7	62.6
Germany	72.8	71.1	71.9	75.8	73.0	74.4	71.2	69.7	70.4
Ghana	52.2	53.0	52.4	54.5	54.1	54.2	51.4	52.0	51.6
Greece	76.5	73.9	75.1	78.8	74.9	76.8	75.2	74.4	74.9
Guatemala	63.6	60.5	62.0	60.6	57.8	58.9	63.2	60.8	61.9
Hungary	72.0	67.2	69.5	73.1	67.5	70.2	71.6	66.6	69.0
India	47.7	50.1	48.9	55.3	55.5	55.3	47.3	50.0	48.8
Ireland	77.7	75.5	76.6	79.1	76.2	77.7	77.0	74.8	75.9
Israel	70.2	70.1	70.2	73.7	72.5	73.1	70.8	71.1	71.0
Italy	73.0	72.1	72.5	76.5	74.7	75.5	71.1	71.2	70.9
Kazakhstan	62.2	55.5	58.5	66.0	57.0	61.2	60.1	54.6	57.2
Kenya	50.1	49.9	50.0	52.4	50.9	51.7	48.9	49.3	49.1
Lao	58.0	58.0	58.0	60.5	59.1	59.8	57.8	56.6	57.2
Latvia	66.3	60.3	62.9	70.3	62.3	66.1	66.5	58.9	62.8
Luxembourg	74.8	72.3	73.7	75.4	71.5	73.7	77.0	74.8	75.9
Malawi	47.6	47.3	47.5	49.4	48.1	48.8	47.3	46.3	46.8
Malaysia	66.9	65.6	66.2	67.4	65.5	66.3	66.8	65.1	65.9

TABLE A. 5.1: Healthy life expectancy by country using three methods (cont.)

	Untruncated linear index (modification 1)			Truncated linear index (WHO)			Untruncated median score (Modifications 1 + 2)		
	females	males	both	females	males	both	females	males	both
Mali	44.2	42.7	43.7	46.3	44.1	45.4	42.6	42.2	42.8
Mauritania	46.9	46.9	46.8	55.9	52.4	54.0	49.0	45.0	46.9
Mauritius	66.3	63.4	64.8	68.4	63.9	66.0	63.5	63.1	63.3
Mexico	68.8	67.2	67.8	71.9	69.2	70.5	68.6	66.6	67.5
Morocco	57.1	55.2	55.8	61.5	54.4	57.5	51.2	54.2	52.7
Myanmar	59.3	55.6	57.3	62.0	56.4	59.2	58.9	54.5	56.6
Nepal	53.8	55.9	54.7	58.4	58.6	58.5	53.3	55.2	54.2
Netherlands	71.5	71.3	70.8	69.8	68.4	68.9	73.3	73.3	72.7
Norway	80.5	77.3	79.0	79.5	77.0	78.5	81.0	77.5	79.4
Pakistan	55.4	56.7	56.1	60.2	59.1	59.5	55.0	57.6	56.8
Paraguay	68.4	65.3	66.6	66.6	62.7	64.4	67.4	66.8	67.0
Philippines	56.0	55.2	55.5	64.9	60.5	62.5	58.5	53.6	55.9
Portugal	72.7	68.1	70.3	76.8	70.2	73.5	69.7	69.8	69.4
Russia	63.1	55.4	58.8	68.6	57.8	62.9	62.2	54.6	58.2
Senegal	43.4	42.0	42.7	48.3	46.9	47.7	37.7	38.8	38.3
Slovakia	70.8	66.0	67.9	74.7	68.2	71.1	67.1	63.0	64.2
Slovenia	71.4	68.5	70.0	73.1	68.6	70.9	70.9	69.1	70.1
SouthAfrica	39.2	40.9	40.1	44.5	44.8	44.7	37.3	39.0	38.1
Spain	77.1	73.2	75.1	80.4	75.2	77.8	75.5	73.3	74.4
SriLanka	67.9	63.7	65.6	73.1	66.9	69.8	66.5	62.7	64.4
Swaziland	33.2	34.5	33.7	38.1	38.5	38.2	32.5	33.6	32.9
Sweden	75.6	75.4	75.4	76.6	75.8	76.1	76.2	73.3	75.1
Tunisia	65.7	64.6	65.0	70.4	67.4	68.7	64.2	64.1	64.0
Turkey	63.1	62.4	62.5	65.4	62.4	63.7	62.0	64.1	62.9
United Arab Emirates	70.8	69.0	69.9	76.5	73.0	74.4	68.7	68.7	68.5
United Kingdom	73.8	72.7	73.1	73.8	72.4	73.0	73.7	72.1	72.8
Ukraine	62.0	55.7	58.5	67.4	57.5	62.2	59.0	55.5	56.8
Uruguay	75.1	70.3	72.8	74.4	69.9	72.2	75.3	69.9	72.7
Vietnam	69.4	67.1	68.2	73.0	68.8	70.8	68.1	65.7	67.0
Zimbabwe	39.8	42.2	40.9	40.4	42.5	41.4	39.5	41.6	40.5

Chapter 6

Incorporating equity-efficiency interactions in cost-effectiveness analysis: three approaches applied to breast cancer control

In collaboration with: Rob Baltussen, Carin Uyl-de Groot, John Bridges and Louis Niessen

Value in Health 13 (5), 573-579

ABSTRACT

Background:

The past decade, medical technology assessment focuses on cost-effectiveness analysis, yet there is an increasing need to consider equity implications of health interventions as well. This paper addresses three equity-efficiency trade-off methods proposed in the literature. Moreover, it demonstrates their impact on cost-effectiveness analyses in current breast cancer control options for women of different age groups.

Methods:

We adapted an existing breast cancer model to estimate cost-effectiveness and equity effects of breast cancer interventions. We applied three methods to quantify the equity-efficiency trade-offs: 1) targeting specific groups, comparing disparities at baseline and in different intervention scenarios; 2) equity weighting, valuing low and high health gains differently; and 3) multi-criteria decision analysis, weighing multiple equity and efficiency criteria. We compared the resulting composite league tables of all approaches.

Results:

The approaches show that a comprehensive breast cancer program, including screening, for women below 75 years of age was most attractive in both the group targeting approach and the equity weighting approach. These method reduced inequality with 56% and at € 1,908 per equity-quality-adjusted life year gained. In the multi-criteria approach the comprehensive program and treatment in stage III breast cancer for women below 75 years of age were most attractive, with both a 0.82 selection probability.

Conclusion:

In the three equity-weighing approaches, targeting women below 75 years of age was more cost-effective and let to more equitable distributions of health. This likely is similar in other fatal diseases with similar age distributions. The approaches may lead to different outcomes in non-fatal diseases.

6.1. INTRODUCTION

The distribution of the disease burden (Eeden et al., 2003; Ward et al., 2004) and treatment benefits (Bradley et al., 2004; Trivedi et al., 2005) are frequently an area of health economics research. In breast cancer control studies reveal differences in disease burden by race (Ward et al., 2004), urbanization (Sheehan et al., 2004), socio-economics status (Lantz et al., 2006) and insurance status (McDavid et al., 2003). These studies typically report disparities in incidence, prevalence, stage distribution, and disease mortality. In some studies differences in quality-adjusted life expectancy are calculated (Bowen et al., 2007). One may distinguish three ways of equity reporting (Lucas, 2002).

Firstly, one may observe differences in health outcomes, such as life expectancy, quality-of-life, and incidence of a condition. Secondly, disparities may be reported in the provision of health care with those with a more severe condition receiving less i.e. vertical equity. Thirdly, inequities may be related to dissimilar use of health care for individuals with the same health i.e. horizontal equity (Lucas, 2002). These three types of equity are interrelated, as utilization of health care is related to health outcomes and both are related to difference in access. In all cases, inequalities may be reduced through the provision of additional health care to underprivileged groups, for example, by differential reimbursement of health packages (McKee and Figueras, 2002).

Descriptive and distinct information about health disparities and cost-effectiveness estimates in relation to health interventions may be available and may give insight. Yet, due to the descriptive nature its use in health policy addressing equity and efficiency, is limited. Such a broad approach to evidence-based priority setting in health programming would use efficiency information on available strategies, as well as their potential for reducing existing disparities. Without this, reduction of inequalities as a policy goal remains a matter of intuition and debate, rather than of systematic evaluation. If so, still, interventions may have differential effects on the distribution of health depending on the way health inequalities are actually defined, measured and addressed.

Methodological studies on the use of equity considerations in cost-effectiveness analysis and its effect on health inequalities are reported (Dolan and Olsen, 2001; Bleichrodt et al., 2004, 2005; Baltussen and Niessen, 2006; Williams and Cookson, 2006). However, comparisons of the impact of these methods in economic evaluation have, so far, not been done and any application in breast cancer control is absent. We distinguished three different methods that can be potentially beneficial in priority setting: targeting specific groups, equity weighting (Bleichrodt et al., 2005) and multi criteria decision analysis (Baltussen and Niessen, 2006; Baltussen et al., 2006).

As the aim of the paper is to show the potential and the impact of these approaches in the use of cost-effectiveness analysis e.g. by government agencies responsible for the selection of health benefit packages. Such processes may have yet to become more explicit, transparent, and thorough if equity implications are to be considered similarly as

and parallel to cost-effectiveness analyses. Our perspective is societal and governmental, given the nature of any operational equity-efficiency approach.

We explain three approaches in section 6.2 and relate them to the underlying theory. Subsequently, we demonstrate their application in cost-effectiveness analyses, aiming at a rank order of optional interventions. We apply the equity-incorporating approaches for breast cancer control evaluations using an existing breast cancer life-table model (Groot et al., 2006), addressing the existing controversy in breast cancer control options by age groups. Differentiating breast cancer control options by age is subject to debate (Boucha-rdy et al., 2003; Holmes and Muss, 2003; Giordano et al., 2005).

6.2. METHODS

We first describe three equitability approaches; subsequently summarize the use and application of the existing WHO breast cancer model and the combined equity and cost-effectiveness analysis.

(a) Equity approaches

(i) Targeting specific groups

The first method we identified for the integration of distributive and economic impact of health interventions is simply targeting specific groups. This method shows how disparities between groups in breast cancer burden can be reduced through interventions in specific population subgroups e.g. on the basis of insurance status (Halpern et al., 2007). The first step in the analysis is identical to the measurement of systematic differences by subgroups, usually defined by an indicator of social economic status. The second step involves selecting an intervention which addresses the difference between those specific subgroups. This means that it is necessary to determine what causes these differences in the first place and in which way they can best be diminished. One or more scenarios can then be constructed in which the targeted group of patients receives the intervention and the remaining group receives the usual level of care. The analyses show the potential improvement in health outcomes of groups of patients both in absolute (increase in health) and relative (reduction of inequalities) terms.

(ii) Equity weighing

Equity weighting (Bleichrodt et al., 2005) is based on the concept of the rank dependent QALY model (Bleichrodt et al., 2004). This method aggregates QALY gains from health interventions over a person's lifetime. In traditional health utility analyses aggregation is straightforward assigning equal valuation to each QALY gained. However, policy makers may want to discriminate between various sub-groups when choosing health

interventions, and may want to give more weight to health gains achieved in relative worse-off groups. Equity weighting quantifies these preferences by assessing the rank of the beneficiaries in the distribution of health. In this approach the valuation of QALYs is non-linear, which makes it possible to assign extra weight to the worst-off. The social value of a QALY profile (i.e. the distribution of health) is then given by:

$$\text{eq 6.1} \quad \sum_{i=1}^{n} \pi_i U(q_i)$$

where π_i is the weight given to the QALY score q for individual i. The non-linearity is shown by the function $U(q_i)$ instead of q_i. In this approach the objective is to maximize eq 6.1 instead of health as described in eq 6.2:

$$\text{eq 6.2} \quad \sum_{i=1}^{n} q_i$$

We assess the values of q_i in different scenarios and compare results with and without equity weighting, using both eq 6.1 and eq 6.2. We use values of equity weights from the Dutch setting (Bleichrodt et al., 2005).

(iii) Multi criteria decision analysis

Multi criteria decision analysis reflects societal preferences for a number of characteristics of health programs in addition to cost effectiveness, such as severity of the disease and the average age of the targeted population (Baltussen and Niessen, 2006). The full set of criteria (Baltussen et al., 2006, 2007) describes the most important aspects of a health intervention. The preferences of society are then measured through a conjoint analysis. First, respondents are simultaneously presented with two health interventions described by the full set of criteria, i.e. a profile. From these two interventions they are asked to pick the most attractive one. A statistical analysis is then used to determine the relative importance of each criterion, reflected by a beta coefficient. Using these coefficients the attractiveness of every intervention described by the full set of criteria can be calculated. The attractiveness of each profile is measured as the probability of selection. Using the probability of selection different interventions can be ranked in a composite league table.

We will use criteria and coefficients from multi criteria investigation among policy makers (Baltussen et al., 2007). The general regression equation is:

$$\text{eq 6.3} \quad P = \frac{EXP(\beta_0 + \sum_{k=1}^{8} \beta_k * X_k)}{1 + EXP(\beta_0 + \sum_{k=1}^{8} \beta_k * X_k)}$$

Where β_0 is a constant, k indicates one of the eight dummy variables (i.e. six criteria of which two have three levels instead of two) and X_k indicates the score of the scenario on the dummy variable (i.e. either zero or one).

As discussed above, all three approaches deal with costs, effects and equity in a different way. Some approaches describe equity implications more thoroughly than others. While some use all cost-effectiveness data and others limit themselves in that regard. Yet all of them ultimately have the same goal: providing an informed equity-efficiency trade-off by ordering health interventions in a composite league table. Table 6.1 summarizes the way the three approaches incorporate the different aspects of health interventions. For illustrative purposes a characterization of traditional health technology assessment is given as well.

TABLE 6.1: Outline different approaches

Approach	Costs	Effect	Equity implication	League table based on:
Traditional	Included	Delta health years	Not applicable	ICER
Targeted intervention	None	Difference in DALE for selected ages per scenario	Absolute and relative reduction of health disparities between groups	Trade-off difference in DALE and costs
Equity weighting	Included	Equity weights * delta health years (i.e. EQALY)	Health benefits weighted more heavily for diseases with larger impact on health	E-ICER
Multi criteria decision analysis	Categorized as an attribute		Equity concept is captured in different attributes	Probabilities

(b) Breast cancer model

We adapted an existing, standardized WHO method to arrive at comparative estimates in a broad cost-effective analysis to a multitude of interventions that is known under the name of the CHOICE program (Tantorres et al., 2003). The existing breast cancer model based on this method (Groot et al., 2006) was adapted in such a way that it is suitable to incorporate the three equity approaches. The breast model distinguishes four breast cancer stages defined according to the guidelines definitions of the American Joint Committee on Cancer (Greene, 2002). Breast cancer stages discern with regard to incidence, prevalence, case-fatality ratio and health state valuation. We used age-specific epidemiological data for the European Region from the Global Burden of Disease Studies (Shibuya et al., 2002). Using age-adjusted data by breast cancer stage (Groot et al., 2006) we computed age and stage specific prevalence, incidence and mortality rates to estimate survival by disease stage. The age-adjusted estimates are provided in Table 6.2.

TABLE 6.2: Model inputs

Disease data		Stage I	Stage II	Stage III	Stage IV	All
Stage distribution at diagnosis	without screening	0.09440	0.14170	0.57970	0.18420	
	with screening	0.49000	0.37440	0.08610	0.04950	
Stage specific mortality	without treatment	0.02000	0.06286	0.15000	0.30000	
	with treatment	0.00638	0.04266	0.09336	0.27500	
Background mortality						0.0102
Prevelence rate						0.0081
Incidence rate						0.0011
Female population						8,002,084
Costs						
Treatment		€ 6,292	€ 6,292	€ 7,066	€ 3,387	
Diagnosis	Breast cancer patient					€ 361
	Healthy woman					€ 86
Follow-up yr 1-5						€ 154
Follow-up yr 6-10						€ 111
Mammography						€ 43
False positive screening	Diagnosis					€ 85
	Biopsy					€ 50

Cost estimates were also derived from the original study by Groot et al. We updated the cost prices to reflect 2007 prices by using a price index of 116.3 (2000 = 100) (OECD. stat, 2008). We included patient costs based on treatment, follow-up, diagnostic work-up of healthy women, screening and false positive cases. Table 6.2 lists associated cost. Furthermore, we assumed equal access to care and equal effectiveness of care for each age group in the existing situation.

First, we constructed a baseline, counterfactual scenario for all patients, irrespective of breast cancer stage or age, in the absence of treatment (Tantorres et al., 2003; Groot et al., 2006). In the intervention scenarios we implemented different sets of treatment in which the stage-specific case-fatalities and/or distribution of incident breast cancer cases were improved for either of the two age groups. The effects of the intervention scenarios were calculated over a standard time period for the whole of the breast cancer population. The intervention scenarios also differed with respect to which patient groups were treated, defined by the stage of breast cancer and age. The groups were analyzed separately and are women aged 75 or less and those over 75. Summarizing, we assessed twelve scenarios: six interventions (the treatment of individual breast cancer stages I, II, III, or IV; the

treatment of all stages; the treatment of all stages plus mammography screening) for two groups of women separately i.e. those below 75 years of age and those of 75 years and over.

We first applied these scenarios in the simple reporting approach to demonstrate the potential for diminishing health inequalities between the two age groups. Secondly, we used equity weighting to revalue the gains of interventions in the two groups. Thirdly, we applied multi criteria decision analysis to breast cancer treatment and screening for the different age groups. Each approach provided a different league table of the investigated health interventions that consider different breast cancer interventions and target groups.

(c) Combined equity and efficiency analysis

(i) Targeting specific groups

We applied the breast cancer model to the EuroA zone of the WHO Global Burden of Disease Regions, commonly used for WHO CHOICE analyses. Analyses were performed separately for women below 75 years and those 75 years and over. As prevalent cases are inherently untreated at t=0 and more numerous in the higher age group, we only include incident cases. We calculated outcomes for both age groups without any treatment. To ensure all health benefits were included, the time horizon was hundred years. To correct for the women's lower life expectancy in higher age groups, we also included past life years, which were assumed to be spent in good health.

In the base case scenario we calculated the average unadjusted and quality-adjusted life years (QALYs) and determined the initial health disparity between both groups. Subsequently, we introduced the six interventions one by one for each group separately and monitored the changes in health disparities. In other words, in each step and for each combination of interventions we computed the difference in health between women below 75 years and those 75 and over.

(ii) Equity weighing

Equity-preference weights were derived from preferences of the general population of the Netherlands by Bleichrodt et al. (Bleichrodt et al., 2005). The weights were differentiated by age and stage of breast cancer at onset. Health effects were first calculated as the number of unweighted health years gained per treated woman and subsequently weighted, applying the weights found in the Bleichrodt study. These weighted health years are the multiplication of the total health effect for breast cancer patients at a certain age with the corresponding equity weight. We refer to this result as equity and quality adjusted life years (EQALYs).

(iii) Multi-criteria decision analysis

In the multi-criteria approach we valued health interventions using the different characteristics of the interventions and the patient groups they are intended for. An index of the characteristics of the breast cancer intervention is used, from published discrete choice experiments (Baltussen et al., 2007). We characterised the breast cancer interventions according to the descriptive criteria used: severity of disease, number of potential beneficiaries, age of target group, individual health benefits, poverty reduction and cost-effectiveness. All of these variables have either two or three ordinal levels. The coefficients for the levels of the attributes were derived from a North American study (Kinter, 2009). Through application of the coefficients for the indexed levels for each intervention in eq 6.3 an overall probability was calculated, reflecting the relative attractiveness of the scenario.

Comparing our findings, we compared all 12 scenarios (6 interventions for 2 different groups) for all three approaches in a new, composite league table in which the rank order was determined by the different outcome measures in each equity approach. Hence, the ranking resulting from the traditional cost-effectiveness analysis was determined by the cost effectiveness ratios (CER); the group targeting ranking depended on the relative reduction in the health gap (largest reduction preferred); the CER based on EQALYs was used to rank the scenarios in the equity weighting approach; and the probability of selection was used to order the results of the multi criteria decision analysis.

6.3. RESULTS

In Table 6.3 the effects of breast cancer intervention on the disparities between the two age groups are presented. In the initial situation, without any breast cancer treatment for either group, the higher age group live 4.87 life time QALYs longer (i.e. the difference in past and future health). Treating the group of breast cancer patients below 75 year of age reduces disparities with a maximum of 56% (screening and treatment of all breast cancer stages); treating the group of women over 80 years of age would increase the difference between the groups up to 6.55 lifetime QALYs, which is an increase of 34.5%.

The results per patient of both the traditional cost-effectiveness analysis and the cost-effectiveness analysis with equity weighting are presented in Table 6.4. The largest costs and effects (both QALYs and EQALYs) are associated with the scenario that includes screening and treatment of all breast cancer stages. The screening and treatment of women below 75 years is the most cost-effective scenario in both the traditional and equity weighting approach. The costs associated with the scenarios differ less than €500 between the two groups of women, but the difference in effects are relatively large in favour of women below 75 years.

TABLE 6.3: Targeting specific groups results

			Treatment women over 75 years						
			None	Stage I	Stage II	Stage III	Stage IV	All stages	All stages after screening
Treatment women under 75 years	None	QALY gap	**-4.87**	-4.91	-4.93	-5.23	-4.89	-5.36	-6.54
		Absolute reduction	**0**	0.04	0.06	0.36	0.02	0.49	1.68
		Relative reduction	**0.00%**	-0.90%	-1.30%	-7.50%	-0.50%	-10.10%	-34.50%
	Stage I	QALY gap	-4.83	-4.88	-4.89	-5.2	-4.86	-5.32	-6.51
		Absolute reduction	-0.03	0.01	0.03	0.33	-0.01	0.46	1.64
		Relative reduction	0.70%	-0.20%	-0.60%	-6.80%	0.20%	-9.40%	-33.80%
	Stage II	QALY gap	-4.81	-4.85	-4.87	-5.17	-4.83	-5.3	-6.49
		Absolute reduction	-0.06	-0.01	0	0.31	-0.03	0.43	1.62
		Relative reduction	1.20%	0.30%	-0.10%	-6.30%	0.70%	-8.90%	-33.30%
	Stage III	QALY gap	-4.44	-4.49	-4.5	-4.81	-4.47	-4.93	-6.12
		Absolute reduction	-0.42	-0.38	-0.36	-0.06	-0.4	0.07	1.25
		Relative reduction	8.70%	7.80%	7.40%	1.30%	8.30%	-1.40%	-25.80%
	Stage IV	QALY gap	-4.82	-4.87	-4.88	-5.19	-4.85	-5.31	-6.5
		Absolute reduction	-0.04	0	0.02	0.32	-0.02	0.45	1.63
		Relative reduction	0.90%	0.00%	-0.40%	-6.60%	0.40%	-9.20%	-33.60%
	All stages	QALY gap	-4.31	-4.35	-4.37	-4.67	-4.33	-4.8	-5.98
		Absolute reduction	-0.56	-0.52	-0.5	-0.2	-0.54	-0.07	1.12
		Relative reduction	11.50%	10.60%	10.20%	4.00%	11.00%	1.40%	-23.00%
	All stages after screening	QALY gap	-2.13	-2.18	-2.2	-2.5	-2.16	-2.63	-3.81
		Absolute reduction	-2.73	-2.69	-2.67	-2.37	-2.71	-2.24	-1.06
		Relative reduction	56.20%	55.30%	54.90%	48.70%	55.70%	46.10%	21.70%

TABLE 6.4: Results usual approach and equity weighting

	Costs		Effects				Cost effectiveness ratio			
			Normal		Equity weighted		Normal		Equity weighted	
Treatment	75-	75+	75-	75+	75-	75+	75-	75+	75-	75+
Stage I	€ 10,130	€ 9,672	1.57	0.376	1.925	0.286	6,454	25,754	5,264	33,862
Stage II	€ 10,044	€ 9,570	1.556	0.37	2.016	0.285	6,456	25,889	4,982	33,564
Stage III	€ 10,777	€ 10,305	2.388	0.556	3.325	0.439	4,513	18,537	3,242	23,498
Stage IV	€ 6,250	€ 5,839	0.389	0.084	0.594	0.069	16,054	69,688	10,529	84,455
All stages	€ 9,779	€ 9,320	1.825	0.426	2.505	0.334	5,358	21,895	3,905	27,867
All stages and screening	€ 16,788	€ 17,006	6.896	1.445	8.801	1.126	2,435	11,766	1,908	15,105

In Table 6.5 all 12 scenarios are characterized using the set of variables described by Baltussen et al. (Baltussen et al., 2007). The resulting probabilities of selection are given in the bottom row. The most probable interventions for selection are the treatment of stage III or the treatment of all stages for women under the age of 75 (P = 0.821) and the least probable is the exclusive treatment of stage I breast cancer for elderly women (P = 0.199). The probabilities show that selecting interventions aimed at the group of women below 75 years of age are considered to be more attractive than those that are aimed at those of 75 and higher.

Table 6.6 summarizes the results of the different interventions in separate composite league tables. Treating women aged below 75 years of age is more attractive than treating women over 75 years of age in all three approaches. In addition, all approaches except the multi criteria decision analysis rank the extensive program (i.e. screening with all types treatment) as most attractive.

6.4. DISCUSSION

Interventions aimed at women less than 75 years of age rank higher in all of the equity-including approaches. Treatment scenarios for women of 75 and over lead to larger health disparities between the two age groups; have lower equity-adjusted cost-effectiveness ratios; and are less likely to have a high probability of selection in applying multiple criteria.

The results of the target group approach can be explained by the relative high average loss of healthy life years at breast cancer diagnoses among ages below 75 in the base case scenario (i.e. no treatment for either group). Women diagnosed with breast cancer aged 80 already have lived 10 more years in good health than women diagnose at age 70. Therefore, any intervention aimed at the disadvantaged group (i.e. women of lower ages with breast cancer) in the calculation will result in positive effects on the distributional of health outcomes across all ages.

The use of equity weights further increases the attractiveness of treating women below 75 years of age and those of 75 and over, as compared to the regular efficiency approach. This is because equity weights are higher for women below 75 years than for women with higher ages, due to the differences in years lived in good health and the remaining potential life span.

The multi-criteria decision analysis shows that having a disease at lower ages is an important criterion. Hence, interventions aimed at women below the age of 75 are more likely to be on the top of the league table. There are other contributing factors as the number of potential beneficiaries is higher, the cost-efficiency ratio is lower, and the net individual health benefits are higher in this age group.

TABLE 6.5: Input values and results multi criteria decision analysis

	Intervention											
Attribute	**Stage I**		**Stage II**		**Stage III**		**Stage IV**		**All stages**		**Screening + all stages**	
	75-	75+	75-	75+	75-	75+	75-	75+	75-	75+	75-	75+
Severity of disease	Not severe		Not severe		Severe		Severe		Severe		Not severe	
Number of potential beneficiaries	Few	Few	Many	Few	Many	Many	Few	Few	Many	Many	Many	Many
Age of target group	Middle-age	Elderly	Middle-age	Elderly	Middle-age	Elderly	Middle-age	Elderly	Middle-age	Elderly	Middle-age	Elderly
Individual health benefits	Large	Small	Large	Small	Large	Small	Small	Small	Large	Small	Large	Large
Poverty reduction	Neutral		Neutral		Neutral		Neutral		Neutral		Neutral	
Cost-effectiveness	CE	CE	CE	CE	CE	CE	CE	Not CE	CE	CE	CE	CE
Probability of selection	0.52	0.509	0.685	0.509	0.821	0.676	0.591	0.199	0.821	0.676	0.685	0.766

TABLE 6.6: League tables for breast cancer interventions in two age groups, for three different equity approaches

	Usual approach		*Targeted intervention*		*Equity weighting*		*Multi criteria decision analysis*	
Scenario	Rank	CER	Rank	Rel. reduc.	Rank	ECER	Rank	Prob. of select.
Stage I for women under 75	9	6,454	6	1%	8	5,264	9	0.520
Stage II for women under 75	8	6,456	4	1%	9	4,982	4	0.685
Stage III for women under 75	11	4,513	3	9%	11	3,242	1	0.821
Stage IV for women under 75	6	16,054	5	1%	7	10,529	8	0.591
All stages for women under 75	10	5,358	2	12%	10	3,905	1	0.821
All stages and screening for women under 75	12	2,435	1	56%	12	1,908	4	0.685
Stage I for women over 75	3	25,754	8	-1%	2	33,862	10	0.509
Stage II for women over 75	2	25,889	9	-1%	3	33,564	10	0.509
Stage III for women over 75	5	18,537	10	-7%	5	23,498	6	0.676
Stage IV for women over 75	1	69,688	7	0%	1	84,455	12	0.199
All stages for women over 75	4	21,895	11	-10%	4	27,867	6	0.676
All stages and screening for women over 75	7	11,766	12	-34%	6	15,105	3	0.766

Although, we adapted the model, distinguishing different breast cancer stages, there are constraints using the present breast cancer model (Groot et al., 2006). Yet we did not incorporate implementation costs. This may vary by group. In addition, the input data used for the analysis (i.e. the epidemiological data, equity weights and the coefficients in the multi-criteria decision analysis) were not all gathered in a similar setting. Consequently the measured equity impact does not reflect true distributional preferences for a single assessed population. For these reasons, our results should be used with caution. However, the significance of this research does not depend on the exact numbers and position in the leagues tables for the selected interventions. Testing and comparing the potential and its order of magnitude of the three existing equity-efficiency trade-off applications in health technology assessment is the main thrust of our paper.

Nevertheless, we consider the preference of targeting women in lower age group as rather robust in the case of breast cancer control. The methods applied this area do not result in large differences between them. Research that addresses interventions for different diseases, among more heterogeneous groups, and distributed differently across age groups, may not show similar patterns and cross-consistency among the three new league tables. This could potentially make the selection of a single equity-efficiency trade-off approach a delicate matter.

We defined equity in health in terms of changes of health outcomes given a particular condition and not in term of health care access or net heath gains. We believe equity

considerations should be concerned with the presence, severity and duration of illness as well as with longevity and life time benefits i.e. the fair innings principle. Hence, health should be measured in terms of disability-adjusted life expectancy (DALE), health-adjusted life expectancy (HALE) or lifetime quality-adjusted life years (QALYs). These measures combine severity of illness and the fair innings principle and may account for the prevailing concepts of equity (Stolk et al., 2005). We believe that the three approaches used in this paper all incorporated (some of) these properties of equity.

However, the approaches do not deal with equity in the same way; they all have their own strengths and weaknesses and seem to be more complementary than mutual exclusive. Simply targeting underprivileged groups is the least comprehensive way to go about the equity-efficiency trade-off as it does not require any prior data collection on preferences of the general public or policy makers. In addition, a health gap is an intuitively understandable measure of differences in health between groups as well as the gap-reducing effect of interventions, either in absolute or relative terms. However, targeting specific groups does not simultaneously assess the impact an intervention has on equity and its cost-effectiveness. This means that the trade-off between equity and efficiency remains implicit, albeit it becomes less transparent.

Equity weighting explicitly combines those preferences with associated costs and health effects in an easily understandable measure: the equity adjusted cost-effectiveness ratio, incorporating the willingness to give up life years for equity reasons. This measure can be interpreted as a regular cost-effectiveness ratio. Although it may be easy to interpret equity weights and equity-weighted outcomes, it may be difficult to understand in which way equity weights are measured and calculated.

Multi-criteria decision analysis incorporates more aspects of health interventions than the other two approaches. However, preferences about the relative importance of different criteria are measured in a different context and may be situation-determined. Another disadvantage of this approach is the large amount of information that is lost as the performance on each criterion is categorized, weighted in a single outcome measure i.e. the probability of selection. This limits the possibility to distinguish between different interventions. The number of potential profiles is especially limited if all investigated interventions are aimed at the same disease and same age group as in our example.

Our study shows that there are some applications of the equity-efficiency trade-off at the disposal of policy makers. These applications are potentially promising, because they may lead to better informed and more transparent reimbursement decisions. Better information can result in a shift from intuition-based policy making to more evidence based policy making. Various high-income health care systems are presently shifting towards a process of intervention assessment and appraisal. Typically, the National Institute for Health and Clinical Excellence (NICE) in the United Kingdom has produced a paper on social values. In general, governments can only accomplish more of health goals if the

consequences of health policies for different goals are known (Health and Excellence, 2008).

However, increased quantification of knowledge on the impact of health interventions will reduce the autonomy of policy makers to decide which reimbursement scheme is most attractive and a priority. This also may result in reluctance from policy makers toward such explicit applications of the equity-efficiency trade-off. Therefore it seems most realistic to use the described incorporation of equitability approaches as a potential transparent support of policy with regard to accounting for the equity impact of health interventions.

Chapter 7

Discussion

In this thesis I have made a contribution to the ever growing literature on socio-economic differences in health. This contribution can be divided into three different parts. First, I added to the empirical knowledge on the SES-health gradient. Second, I discussed methodological issues that can arise during the investigation of socio-economic differences in health and proposed some solutions. And third, I provided insight in the consequences of SES-health inequalities for policy. In this chapter I will separately summarize and discuss these three goals and what contribution each chapter makes.

7.1. EMPIRICAL EVIDENCE ON SOCIO-ECONOMIC HEALTH INEQUALITIES

This thesis shows that inequalities are everywhere, but never the same. I have presented evidence from different settings. I found that inequalities in health exist between SES groups in the Netherlands, China and around the world. The extent of the inequalities, however, varies substantially between settings. In Chapter 5 I found that healthy life expectancy might be 25.8 healthy life years larger in rich countries than in poor countries. In chapters 2, 3 and 4 I show that the differences are on average much smaller between individuals within the same country. The extent of inequalities does not only vary geographically, but also over time as illustrated in Chapter 4.

In Chapter 2 I calculated life cycle profiles of poor health in the Netherlands and assessed the impact of selective dropout (see below). The evidence shows that the prevalence of poor health in the Netherlands is much larger in the lower income quartile than in the upper income quartile. This inequality persists at all ages and for both men and women. The inequality however is not constant with age. For men the difference ranges from 10 to 35 percentage points, for women it ranges from 13 to 22 percentage points. Secondly, I find evidence for selective dropout in the Netherlands. Survival probabilities fall much faster for poor individuals who become ill than for rich individuals who become ill. Becoming ill decreases the one year survival probability of a poor 70 year old man 50% more than for a same age male with a high income.

Chapter 3 provides evidence on inequalities in BMI in China between individuals with different levels of SES during childhood. These results deviate from most findings on SES-health inequalities. Usually pre-dominantly U.S. and European evidence suggest higher BMI levels for lower SES groups. The results in Chapter 3 suggest that the effect of SES, as measured by maternal education, on BMI depends on the age of the respondent. At young ages children from mothers with a low education are heavier than their peers from higher educated mothers. BMI increases with age very strongly, but more so for children from higher educated mothers. Consequently, these children have the most severe weight problems. This rotation of the SES-BMI gradient also persists in adulthood SES, higher educated individuals are lighter at young ages and heavier at old ages. The prevalence levels of overweight and obesity are still low in China, but average BMI continues to rise.

Consequently, inequalities in BMI may become increasingly detrimental to public health and have an impact on socio-economic health inequalities.

In overall health, inequalities in China are more in line with traditional findings of better outcomes. Chapter 4 shows that on average health is better for individuals from richer households, albeit at a decreasing rate at higher incomes. The rise of the Erreygers index from 0.0130 to 0.0408 over the period 1991-2006 illustrates that health is becoming increasingly unequally distributed by income. These increased income related health inequalities (IRHI) can be linked to rapid changes in Chinese economy and society (see below). China seems not only to become a wealthier nation, but also a more divided one.

In Chapter 5 I explore worldwide health inequalities using household survey data. As mentioned before the inequalities are very large. Not only is there a gap in healthy life expectancy between the richest and poorest groups of countries, but a gradient is evident at all levels of income. Inequalities are reflected both in unequal life expectancies and unequal average health related quality of life. So far, estimates may have been underestimated as a result of methodological choices. I showed that arguably better choices can be made that reveal an even bigger disparity in healthy life expectancy between rich and poor countries.

Inequalities between individuals from the fairly egalitarian Netherlands are by far not as large as those between the richest and poorest countries in the world. However, in Chapter 6 I show that even in the Netherlands substantial inequalities in health do exist. What is more, these inequalities can be linked back to reimbursement decisions by the Dutch government. I showed that the decision to exclude elderly women for mammography screening can affect the inequality in healthy life span between both groups with more than one healthy life year. I proposed a number of ways to evaluate the initial inequality and its reduction in different reimbursement scenarios. Although each method yielded a different league table for priority setting, all indicated that from both a cost-effectiveness and an inequality point-of-view priority to younger women was maintainable.

7.2. METHODOLOGICAL ISSUES IN THE MEASUREMENT OF SOCIO-ECONOMIC HEALTH INEQUALITIES

All chapters contribute to the way in which health inequalities could or should be measured. They show how methodological issues can impact health inequality measurement. However, as this thesis has shown, to a certain extent each of these issues can be dealt with. In each chapter I have focused on reducing the impact of a single issue. I also show that to account for these issues a considerable amount of effort is necessary. Considering that this thesis by no means deals with all complexities in the relationship between SES and health, researchers are presented with a daunting task.

(a) Causality

In Chapter 3 I deal with causality in the relationship between SES and health, one of the most researched topics in this literature. To summarize, there can be three explanations for the SES-health gradient. First, SES may impact health because income can be used to invest in health or education can lead to better informed health decisions. Secondly, reverse causation may exists since unhealthy individuals will be less able to earn a high income or invest in their education. Finally, there may be third factors that explain the association between SES and health, such as genes. By measuring SES in childhood I reduce the impact of health on SES, because SES in childhood is primarily determined by parental factors and not by individual ones. Although I find a relationship that is largely attributable to effects of SES on health, this finding does not rule out an additional effect of health on SES or of third factors. Even when more complex designs are used these cannot be excluded.

(b) Time- and age-effects

The estimation of health inequalities is further complicated when a time dimension is included. Neither SES nor health is constant over time and therefore the SES-health gradient will also vary over time. A time dimension can be either measured as a time-, age- or cohort effect, but not all simultaneously. In Chapter 2 I specifically investigated age effects in the life cycle profiles of poor health for different income groups in the Netherlands. As expected, for both groups health is very much related to age, but age effects differ considerably between rich and poor. Health deteriorates relatively fast at young ages for the poor, whereas for the rich health starts to deteriorate at a later age. These differential age effects resulted in health inequalities that diverge until age 50 and start to converge from there until around retirement age. Cohort effects turned out to be less important in the life cycle profiles. I hypothesized that these peculiar profiles were (partly) the result of selective dropout. Selective dropout could be important if poor unhealthy individuals would die or move into a nursing home, while rich unhealthy individuals would not. Although I found evidence for such an effect, its impact on the health profiles was limited.

Time affects the measurement of socio-economic differences in health further. Often health and SES are simultaneously assessed in cross-sectional studies. The effect of SES or health may, however, be delayed. Circumstances early in life may already contribute to the SES-health gradient. On the one hand childhood health impacts later life SES (Oreopoulos et al., 2008), on the other early life SES can have an impact on health later in life (Johnson and Schoeni, 2007). In Chapter 3 I deal with this issue by focusing on socio-economic status during childhood and health later in life. The association between maternal education and adulthood BMI in China confirms that the SES-health gradient is co-determined by early life conditions. This relationship depends on the age at which SES

and health are measured, which again underlines the importance of age in the relationship between SES and health.

The early life determinants of socio-economic differences also illustrate another issue related to time. Even those in low SES or poor health early in life can improve their SES or health or vice versa. Cross-sectional relationships do not account for this variation. The inclusion of early life circumstances is not a solution here, since health and SES may have changed in between more than once. In certain cases the issue is less problematic. Education, for example, is mostly time invariant, as are certain chronic diseases. For income it is more complicated because it requires information at different periods in time and knowledge about the timing of effects. Related to the changes in health and SES is the duration of exposure to low SES and poor health. Individuals that are in poor health earlier are faced with the higher risks of negative SES outcomes (Power et al., 1999).

(c) SES, health and societal change

Both SES and health change over time, are related to each other and other factors. Societal change in that respect can have a large impact on the degree of socio-economic health inequalities. In Chapter 4 I have related such changes to income-related health inequalities (IRHI) in China. China has experienced tremendous economic achievements in the past decade, but this has not been without costs. Good accessible health care facilities have been replaced by more privately funded ones (Tang et al., 2008). Income inequality has risen dramatically (Sutherland and Yao, 2011). Finally, pension systems in China are inadequate at securing above poverty level incomes for many elderly Chinese (Cai et al., 2006; Giles et al., 2010). Using a decomposition method I linked the economic developments in China to a tripling of IRHI between 1991 and 2006. The decomposition revealed that poor income growth for elderly women is the main contributor to this growth in IRHI. This finding is in line with the literature, but still cannot incorporate other socio-economic changes such as access to health care. The importance of societal changes in relationship to the SES-health gradient is further confirmed by the geographic sub-analyses in Chapter 4. In China, large differences in development exist between urban/rural and inland/coastal provinces. Pensions, health insurance, access to healthy food and sanitation are all superior in urban coastal areas compared to the rest of China. This is reflected in the degree of IRHI and the income experience of older women.

(d) The measurement of SES and health

So far I have discussed SES and health as well established and easily measured concepts. Both concepts, however, are multi-dimensional and the selected measure will impact results. In Chapter 2, for example, I use income as an indicator of SES and a self-reported health measure as a health variable. Socio-economic status is also reflected in other variables. For example, education (Chapter 3), occupational status (Winkelmann and

Winkelmann, 1998) and wealth (Bond Huie et al., 2003) are all linked to health. Attempts can be made to capture all relevant variance in SES in a composite measure, as well as one for health. In Chapter 5 I make such an attempt, by combining morbidity and mortality variance in healthy life expectancies. Although this improves the measurement of health, it is still doubtful whether the use of five point response scales captures all relevant information. Nonetheless, this tension between the most accurate measurement possible of a single aspect of the health, or SES for that matter, and the complexity of including all aspects of health, lies beneath all attempts to measure the gradient (Adler et al., 2012). The big advantage of the approach applied in Chapter 5, the cross-country comparability of health estimates may outweigh the drawbacks of not measuring all inter-individual variance in health. Inter-individual variance would be more important if the objective of Chapter 5 would have been to capture within country health inequalities. For between country health inequalities, however, comparability between settings is essential.

7.3. THE CONSEQUENCES OF SOCIO-ECONOMIC HEALTH INEQUALITIES

The reduction of health inequalities in general and socio-economic health inequalities more specifically is a main objective of health care policies on the national and international level. Therefore, in addition to measuring the SES-health gradient as accurately as possible, it is important for researchers to make their results useful for policy. Outcomes of inequality research should possess some properties to be able to support policy. They could help evaluate the efficacy of specific policies through monitoring. They could assist in priority setting before a decision is made by providing evidence prior to implementation. They can inform policy makers best if outcomes are easily interpretable. In Chapter 6 I have presented three methods that exhibit these properties for health care priority setting. The first method, targeting specific groups, provides a way to compare the inequalities that exists between (SES) groups in society with or without reimbursement of specific health care interventions. These outcomes are available and both easily monitored and interpreted. The second method, the weighting of gained quality of life years of health care based on inequality measures, also provides easily interpretable outcomes. This method is more preference based, but is less easily constructed and limits the free balancing of policy outcomes by decision makers. Finally, I show how easily interpretable outcomes that are available before policy implication can be obtained through multi-criteria decision making. This method includes inequality outcomes as one of many relevant decision attributes. This comes at the cost of less careful valuation of inequalities, since interventions are in general only ranked as inequality increasing, neutral or decreasing.

The other chapters of this thesis focus less on evidence based policy making, but do have some policy implications. The life cycle profiles of poor health in the Netherlands analyzed in Chapter 2 suggest that health around retirement age may be a limiting factor

for labor capacity, especially for low paid workers. This finding is not new and is already used as an argument against an increase of the retirement age in the Netherlands. Chapter 2 confirms this finding and adds to the evidence by eliminating a potential bias in the measurement of this finding.

The evidence from chapters 3 and 4 is worrying for the Chinese government. Chapter 3 suggests that weight problems will continue to increase and impact public health. What is more, there is a SES gradient in these problems. It is essential to better understand this gradient, since its rotation with age may imply different underlying explanations. It may be a true effect, but can also be a result of cohort effects, strong variations in age effects or differential SES-health gradients for specific groups, e.g. based on gender or region. Chapter 4 may even be more alarming for a Chinese government that claims to strive for a *harmonious society*. The tripling of the Erreygers index between 1991 and 2006 shows that Chinese society has become much less egalitarian. The decomposition of these changes links this increase to changes in the income distribution. This suggests that the Chinese government should focus on better income protection for retirees, as well as better public health services for the poor. Finally, the stratified analyses of inland/coastal urban/rural areas points to the need to prioritize other inland areas as well as rural coastal areas.

The results of Chapter 5 are not the concern of a single government, but of the whole international community. The inequalities between the richest and poorest countries are so large that the poorest countries will need assistance to promote population health. Such inequalities may be too large to be reduced just be health and health care interventions. Continued economic development might be more effective and reduce both health and income disparities. This also applies to the SES-health gradient in general. An issue that relates to economic and health circumstances should be resolved both through economic and health policies, especially if causality is bidirectional.

References
Summary
Samenvatting
Dankwoord
Curriculum Vitae

REFERENCES

Adler, N., Bush, N.R., and Pantell, M.S. (2012). Rigor, vigor, and the study of health disparities. Proceedings of the National Academy of Sciences *109*, 17154–17159.

Akin, J.S., Dow, W.H., Lance, P.M., and Loh, C.P.A. (2005). Changes in access to health care in China, 1989-1997. Health Policy and Planning *20*, 80–89.

Allanson, P.F., Gerdtham, U.-G., and Petrie, D. (2010). Longitudinal analysis of income-related health inequality. Journal of Health Economics *29*, 78–86.

Anderson, G., and Hussey, P.S. (2001). Comparing Health System Performance In OECD Countries. Health Aff *20*, 219–232.

Bago d'Uva, T., Doorslaer, E. van, Lindeboom, M., and O'Donnell, O. (2008a). Does reporting heterogeneity bias the measurement of health disparities? Health Economics *17*, 351–375.

Bago d'Uva, T., Lindeboom, M., O'Donnell, O., and Doorslaer, E. van (2011). Education-related inequity in healthcare with heterogeneous reporting of health. Journal of the Royal Statistical Society: Series A (Statistics in Society) *174*, 639–664.

Bago d'Uva, T., O'Donnell, O., and Doorslaer, E. van (2008b). Differential health reporting by education level and its impact on the measurement of health inequalities among older Europeans. Int. J. Epidemiol. *37*, 1375–1383.

Baltussen, R., Asbroek, A.H.A. ten, Koolman, X., Shrestha, N., Bhattarai, P., and Niessen, L.W. (2007). Priority setting using multiple criteria: should a lung health programme be implemented in Nepal? Health Policy Plan. *22*, 178–185.

Baltussen, R., and Niessen, L. (2006). Priority setting of health interventions: the need for multi-criteria decision analysis. Cost Eff Resour Alloc *4*, 14.

Baltussen, R., Stolk, E., Chisholm, D., and Aikins, M. (2006). Towards a multi-criteria approach for priority setting: an application to Ghana. Health Economics *15*, 689–696.

Barba, C., Cavalli-Sforza, T., Cutter, J., Darnton-Hill, I., Deurenberg, P., Deurenberg-Yap, M., Gill, T., James, P., Ko, G., Kosulwat, V., et al. (2004). Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. Lancet *363*, 157–163.

Baum, C.L., and Ruhm, C.J. (2009). Age, socioeconomic status and obesity growth. Journal of Health Economics *28*, 635–648.

Beckett, M. (2000). Converging Health Inequalities in Later Life-An Artifact of Mortality Selection? Journal of Health and Social Behavior *41*, 106–119.

Beckett, M.K., and Elliott, M.N. (2001). Mortality and Sample Selection: Reply to Noymer. Journal of Health and Social Behavior *42*, 328–331.

Bell, A.C., Ge, K., and Popkin, B.M. (2001). Weight gain and its predictors in Chinese adults. International Journal of Obesity and Related Metabolic Disorders : Journal of the International Association for the Study of Obesity *25*, 1079–1086.

Benjamin, D., Brandt, L., Giles, J., and Wang, S. (2008). China's great economic transformation. In China's Great Economic Transformation, L. Brandt, and T.G. Rawski, eds. (Cambridge University Press), pp. 729–775.

Black, R.E., Morris, S.S., and Bryce, J. (2003). Where and why are 10 million children dying every year? The Lancet *361*, 2226–2234.

Bleich, S.N., Jarlenski, M.P., Bell, C.N., and LaVeist, T.A. (2012). Health Inequalities: Trends, Progress, and Policy. Annual Review of Public Health *33*, 7–40.

Bleichrodt, H., Diecidue, E., and Quiggin, J. (2004). Equity weights in the allocation of health care: the rank-dependent QALY model. J Health Econ *23*, 157–171.

Bleichrodt, H., Doctor, J., and Stolk, E. (2005). A nonparametric elicitation of the equity-efficiency trade-off in cost-utility analysis. Journal of Health Economics *24*, 655–678.

Boerma, J.T., Mathers, C., and Abou-Zahr, C. (2010). WHO and Global Health Monitoring: The Way Forward. PLoS Med *7*, e1000373.

Bond Huie, S.A., Krueger, P.M., Rogers, R.G., and Hummer, R.A. (2003). Wealth, Race, and Mortality*. Social Science Quarterly *84*, 667–684.

Bongaarts, J. (2006). How Long Will We Live? Population and Development Review *32*, 605–628.

Borghans, L., and Golsteyn, B.H.H. (2006). Time discounting and the body mass index: Evidence from the Netherlands. Economics and Human Biology *4*, 39–61.

Bouchardy, C., Rapiti, E., Fioretta, G., Laissue, P., Neyroud-Caspar, I., Schafer, P., Kurtz, J., Sappino, A.-P., and Vlastos, G. (2003). Undertreatment Strongly Decreases Prognosis of Breast Cancer in Elderly Women. J Clin Oncol *21*, 3580–3587.

Bowen, D., Alfano, C., McGregor, B., Kuniyuki, A., Bernstein, L., Meeske, K., Baumgartner, K., Fetherolf, J., Reeve, B., Smith, A., et al. (2007). Possible socioeconomic and ethnic disparities in quality of life in a cohort of breast cancer survivors. Breast Cancer Research and Treatment *106*, 85–95.

Bradley, E.H., Herrin, J., Wang, Y., McNamara, R.L., Webster, T.R., Magid, D.J., Blaney, M., Peterson, E.D., Canto, J.G., Jr, C.V.P., et al. (2004). Racial and ethnic differences in time to acute reperfusion therapy for patients hospitalized with myocardial infarction. Journal of the American Medical Association *292*, 1563–1572.

Bramall, C. (2009). Chinese Economic Development (London, New York: Routledge).

Braveman, P., and Gruskin, S. (2003). Defining equity in health. J Epidemiol Community Health *57*, 254–258.

Brun, J.F., Combes, J.-L.C., and Renard, M.-F. (2002). Are there spillover effects between coastal and noncoastal regions in China? China Economic Review *13*, 161–169.

Burkhauser, R.V., and Cawley, J. (2008). Beyond BMI: The value of more accurate measures of fatness and obesity in social science research. Journal of Health Economics *27*, 519–529.

Cai, F., Giles, J., and Meng, X. (2006). How well do children insure parents against low retirement income? An analysis using survey data from urban China. Journal of Public Economics *90*, 2229–2255.

Carolina Population Center (2010). China Health and Nutrition Survey.

Case, A., and Deaton, A. (2005). Broken down by work and sex: How our health declines. Analyses in the Economics of Aging 185–205.

Case, A., Fertig, A., and Paxson, C. (2005). The lasting impact of childhood health and circumstance. Journal of Health Economics *24*, 365–389.

Chang, G.H. (2002). The cause and cure of China's widening income disparity. China Economic Review *13*, 335–340.

Chen, J., and Fleisher, B.M. (1996). Regional income inequality and economic growth in China. Journal of Comparative Economics *22*, 141–164.

Chen, Z., Eastwood, D.B., and Yen, S.T. (2007). A decade's story of childhood malnutrition inequality in China: Where you live does matter. China Economic Review *18*, 139–154.

Chen, Z., and Meltzer, D. (2008). Beefing up with the Chans: Evidence for the effects of relative income and income inequality on health from the China Health and Nutrition Survey. Social Science & Medicine *66*, 2206–2217.

Colombo, F., Llena-Nozal, A., Mercier, J., and Tjadens, F. (2011). Help Wanted? Providing and Paying for Long-Term Care: Providing and Paying for Long-Term Care (OECD Publishing).

Commission on Social Determinants of Health (CSDH) (2008). Closing the gap in a generation: health equity through action on the social determinants of health. Final Report of the Commission on Social Determinants of Health. (Geneva: World Health Organization).

Contoyannis, P., and Forster, M. (1999). The distribution of health and income: a theoretical framework. Journal of Health Economics *18*, 603–620.

Cui, Z., and Dibley, M.J. (2012). Trends in Dietary Energy, Fat, Carbohydrate and Protein Intake in Chinese Children and Adolescents from 1991 to 2009. British Journal of Nutrition *FirstView*, 1–8.

Currie, A., Shields, M.A., and Price, S.W. (2007). The child health/family income gradient: Evidence from England. Journal of Health Economics *26*, 213–232.

Currie, J. (2009). Healthy, Wealthy, and Wise: Socioeconomic Status, Poor Health in Childhood, and Human Capital Development. Journal of Economic Literature *47*, 87–122.

Cutler, D., Deaton, A., and Lleras-Muney, A. (2006). The Determinants of Mortality. The Journal of Economic Perspectives *20*, 97–120.

Cutler, D.M., Glaeser, E.L., and Shapiro, J.M. (2003). Why Have Americans Become More Obese? The Journal of Economic Perspectives *17*, 93–118.

Deaton, A., and Paxson, C. (1998a). Health, Income, and Inequality over the Life Cycle. (National Bureau of Economic Research, Inc), pp. 431–462.

Deaton, A.S., and Paxson, C.H. (1998b). Aging and Inequality in Income and Health. The American Economic Review *88*, 248–253.

Department of Health (2003). Tackling Health Inequalities: A Programme for Action (London: Department of Health).

Deurenberg-Yap, M., and Deurenberg, P. (2003). Is a Re-evaluation of WHO Body Mass Index Cut-off Values Needed? The Case of Asians in Singapore. Nutrition Reviews *61*, S80–S87.

Dolan, P. (1997). Modeling Valuations for EuroQol Health States. Medical Care *35*, 1095.

Dolan, P., and Olsen, J.A. (2001). Equity in health: the importance of different health streams. Journal of Health Economics *20*, 823–834.

Doorslaer, E. van, and Jones, A.M. (2003). Inequalities in self-reported health: validation of a new approach to measurement. Journal of Health Economics *22*, 61–87.

Doorslaer, E. van, and Koolman, X. (2004). Explaining the differences in income-related health inequalities across European countries. Health Economics *13*, 609–628.

Eeden, S.K.V.D., Tanner, C.M., Bernstein, A.L., Fross, R.D., Leimpeter, A., Bloch, D.A., and Nelson, L.M. (2003). Incidence of Parkinson's disease: Variation by age, gender, and race/ethnicity. American Journal of Epidemiology *157*, 1015–1022.

Erreygers, G. (2009a). Correcting the Concentration Index. Journal of Health Economics *28*, 504–515.

Erreygers, G. (2009b). Correcting the Concentration Index: A reply to Wagstaff. Journal of Health Economics *28*, 521–524.

Erreygers, G., and Van Ourti, T. (2011a). Measuring socioeconomic inequality in health, health care and health financing by means of rank-dependent indices: A recipe for good practice. Journal of Health Economics *30*, 685–694.

Erreygers, G., and Van Ourti, T. (2011b). Putting the cart before the horse. A comment on Wagstaff on inequality measurement in the presence of binary variables. Health Economics *20*, 1161–1165.

Etilé, F., and Milcent, C. (2006). Income-related reporting heterogeneity in self-assessed health: evidence from France. Health Economics *15*, 965–981.

Ettner, S.L. (1996). New evidence on the relationship between income and health. Journal of Health Economics *15*, 67–85.

Feldstein, M. (1999). Social security pension reform in China. China Economic Review *10*, 99–107.

Feng, X.L., Zhu, J., Zhang, L., Song, L., Hipgrave, D.B., Guo, S., Ronsmans, C., Guo, Y., and Yang, Q. (2010). Socio-economic disparities in maternal mortality in China between 1996 and 2006. British Journal of Obstetrics and Gynaecology *117*, 1527–1536.

Fortin, N., Lemieux, T., and Firpo, S. (2011). Decomposition Methods in Economics. In Handbook of Labor Economics, A. Orley, and C. David, eds. (Amsterdam: Elsevier), pp. 1–102.

Fries, J.F. (2000). Compression of morbidity in the elderly. Vaccine *18*, 1584–1589.

García-Gómez, P., Kippersluis, H. van, O'Donnell, O., and Doorslaer, E. van (2013). Effects of health on own and spousal employment and income using acute hospital admissions. Journal of Human Resources.

Giles, J., Wang, D., and Zhao, C. (2010). Can China's Rural Elderly Count on Support from Adult Children? Implications of Rural-to-Urban Migration. Journal of Population Ageing *3*, 183–204.

Giordano, S.H., Hortobagyi, G.N., Kau, S.-W.C., Theriault, R.L., and Bondy, M.L. (2005). Breast Cancer Treatment Guidelines in Older Women. J Clin Oncol *23*, 783–791.

Graham, H., and Kelly, M.P. (2004). Health inequalities : concepts, frameworks and policy (London: Health Development Agency).

Greene, F.L. (2002). AJCC cancer staging manual (New York: Springer-Verlag).

Groot, M.T., Baltussen, R., Groot, C.A.U., Anderson, B.O., and Hortobagyi, G.N. (2006). Costs and health effects of breast cancer interventions in epidemiologically different regions of Africa, North America, and Asia. Breast J *12 Suppl 1*, S81–90.

Groot, W. (2000). Adaptation and scale of reference bias in self-assessments of quality of life. Journal of Health Economics *19*, 403–420.

Gustafsson, B., Shi, L., and Sicular, T. (2008). Inequality and Public Policy in China (Cambridge University Press).

Gustafsson, B., Shi, L., and Zhong, W. (2006). The distribution of wealth in urban China and in China as a whole in 1995. Review of Income and Wealth *52*, 173–188.

Gutierrez-Fisac, J.L., Guallar-Castillon, P., Diez-Ganan, L., Garcia, E.L., Banegas, J.R.B., and Artalejo, F.R. (2002). Work-Related Physical Activity Is Not Associated with Body Mass Index and Obesity. Obesity *10*, 270–276.

Gwatkin, D.R., Rutstein, S., Johnson, K., Suliman, E., Wagstaff, A., and Amouzou, A. (2007). Socio-economic differences in health, nutrition, and population within developing countries: an overview (Washington, DC: The World Bank).

Halpern, M.T., Bian, J., Ward, E.M., Schrag, N.M., and Chen, A.Y. (2007). Insurance status and stage of cancer at diagnosis among women with breast cancer. Cancer *110*, 403–411.

Harris, J.A., and Benedict, F.G. (1919). A Biometric Study of the Basal Metabolism in Man (Carnegie Institution of Washington).

Health, N.I. for, and Excellence, C. (2008). Social value judgements: principles for the development of NICE guidance.

Herd, P. (2006). Do Functional Health Inequalities Decrease in Old Age?: Educational Status and Functional Decline Among the 1931-1941 Birth Cohort. Research on Aging *28*, 375–392.

Holmes, C.E., and Muss, H.B. (2003). Diagnosis and Treatment of Breast Cancer in the Elderly. CA Cancer J Clin *53*, 227–244.

Hougaard, J.L., Østerdal, L.P., and Yu, Y. (2011). The Chinese healthcare system: structure, problems and challenges. Applied Health Economics and Health Policy *9*, 1–13.

House, J.S., Lantz, P.M., and Herd, P. (2005). Continuity and Change in the Social Stratification of Aging and Health Over the Life Course: Evidence From a Nationally Representative Longitudinal Study From 1986 to 2001/2002 (Americans' Changing Lives Study). J Gerontol B Psychol Sci Soc Sci *60*, S15–26.

Hu, G., Hu, G., Pekkarinen, H., Hänninen, O., Tian, H., and Jin, R. (2002). Comparison of dietary and non-dietary risk factors in overweight and normal-weight Chinese adults. Br. J. Nutr. *88*, 91–97.

Idler, E.L., and Benyamini, Y. (1997). Self-rated health and mortality: a review of twenty-seven community studies. J Health Soc Behav *38*, 21–37.

Idler, E.L., and Kasl, S.V. (1995). Self-Ratings of Health: Do they also Predict change in Functional Ability? The Journals of Gerontology Series B: Psychological Sciences and Social Sciences *50B*, S344–S353.

Johnson, R.C., and Schoeni, R.F. (2007). The Influence of Early-Life Events on Human Capital, Health Status, and Labor Market Outcomes Over the Life Course.

Jones, A.M., Koolman, X., and Rice, N. (2006). Health-related non-response in the British Household Panel Survey and European Community Household Panel: using inverse-probability-weighted estimators in non-linear models. Journal of the Royal Statistical Society: Series A (Statistics in Society) *169*, 543–569.

Jones, A.M., and Nicolás, A.L. (2004). Measurement and explanation of socioeconomic inequality in health with longitudinal data. Health Economics *13*, 1015–1030.

Jones-Smith, J.C., Gordon-Larsen, P., Siddiqi, A., and Popkin, B.M. (2011). Is the burden of overweight shifting to the poor across the globe|[quest]| Time trends among women in 39 low- and middle-income countries (1991|[ndash]|2008). International Journal of Obesity.

Kalache, A., Aboderin, I., and Hoskins, I. (2002). Compression of morbidity and active ageing: key priorities for public health policy in the 21st century. Bulletin of the World Health Organization *80*, 243–244.

Kanbur, R., and Zhang, X. (2005). Fifty Years of Regional Inequality in China: a Journey Through Central Planning, Reform, and Openness. Review of Development Economics *9*, 87–106.

Kant, A.K., and Graubard, B.I. (2005). Energy density of diets reported by American adults: association with food group intake, nutrient intake, and body weight. Int J Obes Relat Metab Disord *29*, 950–956.

Kapteyn, A., Smith, J.P., and Soest, A. van (2009). Work Disability, Work, and Justification Bias in Europe and the U.S. (National Bureau of Economic Research).

Keidel, A.I. (2009). Chinese regional inequalities in income and well-being. Review of Income and Wealth *55*, 538–561.

Kelly, T., Yang, W., Chen, C.-S., Reynolds, K., and He, J. (2008). Global burden of obesity in 2005 and projections to 2030. International Journal of Obesity *32*, 1431–1437.

Kerkhofs, M., and Lindeboom, M. (1995). Subjective health measures and state dependent reporting errors. Health Economics *4*, 221–235.

Khor, N., and Pencavel, J.H. (2010). Poverty, Inequality, and Inclusive Growth in Asia: Measurement, Policy Issues, and Country Studies. In Evolution of Income Mobility in the People's Republic of China: 1991-2002, J. Zhuang, ed. pp. 387–408.

Kim, S., Symons, M., and Popkin, B.M. (2004). Contrasting Socioeconomic Profiles Related to Healthier Lifestyles in China and the United States. Am. J. Epidemiol. *159*, 184–191.

Kind, P., Dolan, P., Gudex, C., and Williams, A. (1998). Variations in population health status: results from a United Kingdom national questionnaire survey. BMJ *316*, 736–741.

King, G., Murray, C.J.L., Salomon, J.A., and Tandon, A. (2003). Enhancing the Validity and Cross-Cultural Comparability of Measurement in Survey Research. The American Political Science Review *97*, 567–583.

King, H., Aubert, R.E., and Herman, W.H. (1998). Global Burden of Diabetes, 1995–2025: Prevalence, numerical estimates, and projections. Dia Care *21*, 1414–1431.

Kinter, E. (2009). Widening the scope of comparative effectiveness analysis in reimbursement decisions – surveys of policy-makers in Europe, the US and Canada. In iHEA 7th World Congres: Harmonizing Health and Economics,.

Kippersluis, H. van, O'Donnell, O., and Doorslaer, E. van (2009a). Long Run Returns to Education: Does Schooling Lead to an Extended Old Age? The Journal of Human Resources *4*, 1.

Kippersluis, H. van, O'Donnell, O., Doorslaer, E. van, and Van Ourti, T. (2010). Socioeconomic differences in health over the life cycle in an Egalitarian country. Social Science & Medicine *70*, 428–438.

Kippersluis, H. van, Van Ourti, T., O'Donnell, O., and Doorslaer, E. van (2009b). Health and income across the life cycle and generations in Europe. Journal of Health Economics *28*, 818–830.

Knight, J., Shi, L., and Renwei, Z. (2004). Divergent Means and Convergent Inequality of Incomes among the Provinces and Cities of Urban China (World Institute for Development Economic Research (UNU-WIDER)).

Kuczmarski, M.F., Kuczmarski, R.J., and Najjar, M. (2001). Effects of age on validity of self-reported height, weight, and body mass index: Findings from the third National Health and Nutrition Examination Survey, 1988-1994. Journal of the American Dietetic Association *101*, 28–34.

Lambert, P.J. (2001). The distribution and redistribution of income (Manchester University Press).

Lambert, P.J., and Zheng, B. (2011). On the consistent measurement of attainment and shortfall inequality. Journal of Health Economics *30*, 214–219.

Lantz, P.M., Mujahid, M., Schwartz, K., Janz, N.K., Fagerlin, A., Salem, B., Liu, L., Deapen, D., and Katz, S.J. (2006). The influence of race, ethnicity, and individual socioeconomic factors on breast cancer stage at diagnosis. Am J Public Health *96*, 2173–2178.

LaVeist, T.A., Gaskin, D.J., and Richard, P. (2009). The Economic Burden of Health Inequalities in the United States.

Lerman, R.I., and Yitzhaki, S. (1989). Improving the accuracy of estimates of Gini coefficients. Journal of Econometrics *42*, 43–47.

Lin, T., Zhuang, J., Yarcia, D., and Lin, F. (2008). Income Inequality in the People's Republic of China and Its Decomposition: 1990-2004. Asian Development Review *25*, 119–136.

Lindeboom, M., and Doorslaer, E. van (2004). Cut-point shift and index shift in self-reported health. Journal of Health Economics *23*, 1083–1099.

Liu, G.G., Dow, W.H., Fu, A.Z., Akin, J., and Lance, P. (2008). Income productivity in China: On the role of health. Journal of Health Economics *27*, 27–44.

Looper, M. de, and Lafortune, G. (2009). Measuring Disparities in Health Status and in Access and Use of Health Care in OECD Countries (OECD Publishing).

Lopez, A.D., and Murray, C.C.J.L. (1998). The global burden of disease, 1990-2020. Nat Med *4*, 1241–1243.

Lu, Y., and Goldman, D. (2010). The Effects of Relative Food Prices on Obesity â€" Evidence from China: 1991-2006 (National Bureau of Economic Research, Inc).

Lucas, A.O. (2002). Health policies in developing countries. R. Detels, J. McEwen, R. Beaglehole, and H. Tanaka, eds. (New York: Oxford University Press), pp. 281–296.

Luo, Z., Mu, R., and Zhang, X. (2006). Famine and Overweight in China. Applied Economic Perspectives and Policy *28*, 296–304.

Lynch, S.M. (2003). Cohort and Life-Course Patterns in the Relationship between Education and Health: A Hierarchical Approach. Demography *40*, 309–331.

Ma, B. (2010). Socioeconomic Status and Overweight and Obesity: How Does the Gradient Change with Age? New Evidence from China. University of Maryland Baltimore County.

Macintyre, S. (2003). Evidence based policy making. BMJ *326*, 5–6.

Mackenbach, J.P. (2002). A strategy for tackling health inequalities in the Netherlands. BMJ *325*, 1029–1032.

Mathers, C., Iburg, K., Salomon, J., Tandon, A., Chatterji, S., Ustun, B., and Murray, C. (2004). Global patterns of healthy life expectancy in the year 2002. BMC Public Health *4*, 66.

Mathers, C.D., Murray, C.J.L., and Salomon, J. (2003a). Methods for measuring healthy life expectancy. Health Systems Performance Assessment: Debates, Methods and Empiricism.

Mathers, C.D., Murray, C.J.L., Salomon, J.A., Sadana, R., Tandon, A., and Lopez, A.D. (2003b). Healthy life expectancy: comparison of OECD countries in 2001. Australian and New Zealand Journal of Public Health *27*, 5–11.

Mathers, C.D., Sadana, R., Salomon, J.A., Murray, C.J.L., and Lopez, A.D. (2001a). Healthy life expectancy in 191 countries, 1999. The Lancet *357*, 1685–1691.

Mathers, C.D., Salomon, J., Murray, C.J.L., and Lopez, A. (2003c). Alternative summary measures of average population health. Health Systems Performance Assessment: Debates, Methods and Empiricism.

Mathers, C.D., Vos, T., Lopez, A.D., Salomon, J., and Ezzati, M. (2001b). National Burden of Disease Studies: A Practical Guide. Edition 2.0. 2001. World Health Organization, Global Program on Evidence for Health Policy, Geneva.

McDavid, K., Tucker, T.C., Sloggett, A., and Coleman, M.P. (2003). Cancer Survival in Kentucky and Health Insurance Coverage. Arch Intern Med *163*, 2135–2144.

McKee, M., and Figueras, J. (2002). Strategies for health services. R. Detels, J. McEwen, R. Beaglehole, and H. Tanaka, eds. (New York: Oxford University Press), pp. 1889–1911.

McLaren, L. (2007). Socioeconomic status and obesity. Epidemiologic Reviews *29*, 29–48.

McMunn, A., Nazroo, J., and Breeze, E. (2009). Inequalities in health at older ages: a longitudinal investigation of the onset of illness and survival effects in England. Age Ageing *38*, 181–187.

Meer, J., Miller, D.L., and Rosen, H.S. (2003). Exploring the health-wealth nexus. Journal of Health Economics *22*, 713–730.

Meng, X., Gregory, R., and Wan, G. (2007). Urban Poverty in China and its contributing factors, 1986–2000. Review of Income and Wealth *53*, 167–189.

Mills, J.A., and Zandvakili, S. (1997). Statistical inference via bootstrapping for measures of inequality. Journal of Applied Econometrics *12*, 133–150.

Minicuci, N., Noale, M., Pluijm, S.M.F., Zunzunegui, M.V., Blumstein, T., Deeg, D.J.H., Bardage, C., Jylhä, M., and the, C. working group for (2004). Disability-free life expectancy: a cross-national comparison of six longitudinal studies on aging. The CLESA project. European Journal of Ageing *1*, 37–44.

Mirowsky, J., and Ross, C.E. (2008). Education and Self-Rated Health: Cumulative Advantage and Its Rising Importance. Research on Aging *30*, 93–122.

Monda, K.L., Adair, L.S., Zhai, F., and Popkin, B.M. (2007). Longitudinal relationships between occupational and domestic physical activity patterns and body weight in China. Eur J Clin Nutr *62*, 1318–1325.

Murray, C.J., and Lopez, A.D. (1997). Regional patterns of disability-free life expectancy and disability-adjusted life expectancy: Global Burden of Disease Study. The Lancet *349*, 1347–1352.

Murray, C.J.L., and Evans, D.B. (2003a). Health Systems Peformance Assessment: Debates, Methods and Empiricism (Geneva: World Health Organization).

Murray, C.J.L., and Evans, D.B. (2003b). Health Systems Perfomance Assessment: Goals, Framework and Overview. J.L. Murray, and D.B. Evans, eds. (Geneva: World Health Organisation),.

MVH Group (1995). The Measurement and Valuation of Health. Final report on the modeling of valuation tariffs (York Centre for Health Economics 1995.).

Ng, S.W., Norton, E.C., and Popkin, B.M. (2009). Why have physical activity levels declined among Chinese adults? Findings from the 1991â€“2006 China health and nutrition surveys. Social Science & Medicine *68*, 1305–1314.

Ng, S.W., Zhai, F., and Popkin, B.M. (2008). Impacts of China's edible oil pricing policy on nutrition. Social Science & Medicine *66*, 414–426.

Noymer, A. (2001). Mortality Selection and Sample Selection: A Comment on Beckett. Journal of Health and Social Behavior *42*, 326–327.

O'Donnell, O., Doorslaer, E. van, Wagstaff, A., and Lindelow, M. (2008). Analyzing Health Equity Using Household Survey Data (Washington, DC: The World Bank).

OECD.stat (2008). Consumer price indices (MEI).

Oreopoulos, P. (2006). Estimating Average and Local Average Treatment Effects of Education When Compulsory Schooling Laws Really Matter. The American Economic Review *96*, 152–175.

Oreopoulos, P., Stabile, M., Walld, R., and Roos, L.L. (2008). Short-, Medium-, and Long-Term Consequences of Poor Infant Health An Analysis Using Siblings and Twins. J. Human Resources *43*, 88–138.

Van Ourti, T., Doorslaer, E. van, and Koolman, X. (2009). The effect of income growth and inequality on health inequality: Theory and empirical evidence from the European Panel. Journal of Health Economics *28*, 525–539.

Petrie, D.J., Allanson, P.F., and Gerdtham, U.-G. (2011). Accounting for the dead in the longitudinal analysis of income-related health inequalities. Journal of Health Economics *30*, 1113–1123.

Philipson, T.J., and Posner, R.A. (2003). The long-run growth in obesity as a function of technological change. Perspect. Biol. Med. *46*, S87–107.

Van de Poel, E., O'Donnell, O., and Doorslaer, E. van (2009). Urbanization and the spread of diseases of affluence in China. Economics & Human Biology *7*, 200–216.

Van de Poel, E., O'Donnell, O., and Doorslaer, E. van (2012). Is there a health penalty of China's rapid urbanization? Health Economics *21*, 367–385.

Power, C., Manor, O., and Matthews, S. (1999). The duration and timing of exposure: effects of socioeconomic environment on adult health. Am J Public Health *89*, 1059–1065.

Preston, S.H. (1975). The Changing Relation between Mortality and Level of Economic Development. Population Studies *29*, 231–248.

Prus, S.G. (2007). Age, SES, and health: a population level analysis of health inequalities over the lifecourse. Sociology of Health & Illness *29*, 275–296.

Sacker, A., Clarke, P., Wiggins, R.D., and Bartley, M. (2005). Social dynamics of health inequalities: a growth curve analysis of aging and self assessed health in the British household panel survey 1991â€"2001. Journal of Epidemiology and Community Health *59*, 495–501.

Sadana, R., Mathers, C.D., Lopez, A.D., Murray, C.J.L., and Iburg, K.M. (2002). Comparative analyses of more than 50 household surveys on health status (Geneva: World Health Organization).

Salomon, A., Mathers, C.D., Chatterji, S., Sadana, R., Üstün, T.B., and Murray, C.J.L. (2003a). Quantifying individual levels of health: Definitions, concepts, and measurement issues. C.J.L. Murray, and D.B. Evans, eds. (Geneva: World Health Organization), pp. 301–318.

Salomon, J.A., Murray, C.J.L., Ustün, B.L., and Chatterji, S. (2003b). Health State Valuations in Summary Measures of Population Health. In Health Systems Performance Assessment: Debates, Methods and Empiricism, C.J.L. Murray, and D.B. Evans, eds. (Geneva: World Health Organization), pp. 409–436.

Salomon, J.A., Tandon, A., Murray, C.J., and World Health Survey Pilot Study Collaborating Groups (2003c). Unpacking Health Perceptions Using Anchoring Vignettes. In Health Systems Performance Assessment: Debates, Methods and Empiricism, J.L. Murray, and D.B. Evans, eds. (Geneva: World Health Organization), pp. 401–408.

Shankar, B. (2009). Obesity in China: The Differential Impacts of Covariates Along the BMI Distribution. Obesity *18*, 1660–1666.

Sheehan, J.T., DeChello, L., Kulldorff, M., Gregorio, D., Gershman, S., and Mroszczyk, M. (2004). The geographic distribution of breast cancer incidence in Massachusetts 1988 to 1997, adjusted for covariates. International Journal of Health Geographics *3*, 17.

Shen, Y., and Yao, Y. (2008). Does grassroots democracy reduce income inequality in China? Journal of Public Economics *92*, 2182–2198.

Shi, J., Liu, M., Zhang, Q., Lu, M., and Quan, H. (2008). Male and Female Adult Population Health Status in China: A Cross-Sectional National Survey. BMC Public Health *8*, 277.

Shibuya, K., Mathers, C., Boschi-Pinto, C., Lopez, A., and Murray, C. (2002). Global and regional estimates of cancer mortality and incidence by site: II. results for the global burden of disease 2000. BMC Cancer *2*, 37.

Sicular, T., Ximing, Y., Gustafsson, B., and Shi, L. (2007). The urban-rural income gap and inequality in China. Review of Income and Wealth *53*, 93–126.

Siermann, C., Teeffelen, P., and Urlings, L. (2004). Equivalentiefactoren 1995-2000 (Voorburg/ Heerlen: Centraal Bureau voor de Statistiek).

Smith, J.P. (2004). Unraveling the SES: Health Connection. In Aging, Health, and Public Policy: Demographic and Economic Perspectives, a Supplement to Population and Development Review 30 *30*, 108–132.

Smith, J.P. (2005). The impact of SES on health over the life course. RAND Working Paper WR-318.

Sobal, J., Hanson, K.L., and Frongillo, E.A. (2009). Gender, Ethnicity, Marital Status, and Body Weight in the United States. Obesity *17*, 2223–2231.

Statistics Netherlands (2011). Statline database.

Stolk, E.A., Pickee, S.J., Ament, A.H., and Busschbach, J.J. (2005). Equity in health care prioritisation: an empirical inquiry into social value. Health Policy *74*, 343–355.

Sturm, R., and Datar, A. (2005). Body mass index in elementary school children, metropolitan area food prices and food outlet density. Public Health *119*, 1059–1068.

Sullivan, D.F. (1988). A single index of mortality and morbidity. The Challenge of Epidemiology: Issues and Selected Readings *86*, 347–354.

Sutherland, D., and Yao, S. (2011). Income Inequality in China Over 30 Years of Reforms. Cambridge J Regions Econ Soc *4*, 91–105.

Tafreschi, D. (2011). The Income Body Weight Gradients in the Developing Economy of China (University of St. Gallen, School of Economics and Political Science).

Tandon, A., Murray, C.J.L., Salomon, J.A., and King, G. (2003). Statiscal Models for Enhancing Cross-Population Comparability. In Health Systems Performance Assessment: Debates, Methods and Empiricism, C.J.L. Murray, and D.B. Evans, eds. (Geneva: World Health Organization), pp. 727–746.

Tang, S., Meng, Q., Chen, L., Bekedam, H., Evans, T., and Whitehead, M. (2008). Tackling the challenges to health equity in China. The Lancet *372*, 1493–1501.

Tantorres, T., Baltussen, R., Adam, T., Hutubessy, R., Acharya, A., Evans, D.B., and Murray, C.J.L. (2003). WHO guide to cost-effectiveness analysis.

Trivedi, A.N., Zaslavsky, A.M., Schneider, E.C., and Ayanian, J.Z. (2005). Trends in the quality of care and racial disparities in medicare managed care. New England Journal of Medicine *353*, 692–700.

U.S. Department of Health and Human Services (2010). Healthy People 2020.

Üstün, T.B., Chatterji, S., Murray, C.J.L., Mechbal, A., and WHS Collaborating Groups (2003a). The World Health Surveys. In Health Systems Performance Assessment: Debates, Methods and Empiricism, C.J.L. Murray, and D.B. Evans, eds. (Geneva.: World Health Organization), pp. 797–808.

Üstün, T.B., Chatterji, S., Villanueva, M., Bendib, L., Çelik, C., Sadana, R., Valentine, N.B., Ortiz, J.P., Tandon, A., Salomon, J.A., et al. (2003b). WHO multi-country survey study on health and responsiveness, 2000–2001. C.J.L. Murray, and D.B. Evans, eds. (Geneva.: World Health Organization), pp. 761–796.

Üstün, T.B., Chatterji, S., Villanueva, M., Bendib, L., Çelik, C., Sadana, R., Valentine, N.B., Ortiz, J.P., Tandon, A., Salomon, J.A., et al. (2003c). WHO multi-country survey study on health and responsiveness, 2000–2001. In Health Systems Performance Assessment: Debates, Methods and Empiricism, C.J.L. Murray, and D.B. Evans, eds. (Geneva.: World Health Organization), pp. 761–796.

Wagstaff, A. (2009). Correcting the concentration index: A comment. Journal of Health Economics *28*, 516–520.

Wagstaff, A. (2011a). Reply to Guido Erreygers and Tom Van Ourti's comment on "The concentration index of a binary outcome revisited."Health Economics *20*, 1166–1168.

Wagstaff, A. (2011b). The concentration index of a binary outcome revisited. Health Economics *20*, 1155–1160.

Wagstaff, A., and Doorslaer, E. van (2000). Income Inequality and Health: What Does the Literature Tell Us? Appl Health Econ Health Policy *21*, 543–567.

Wagstaff, A., Doorslaer, E. van, and Watanabe, N. (2003). On decomposing the causes of health sector inequalities with an application to malnutrition inequalities in Vietnam. Journal of Econometrics *112*, 207–223.

Wagstaff, A., Lindelow, M., Wang, S., and Zhang, S. (2009a). Reforming China's Rural Health System (Washington DC: World Bank).

Wagstaff, A., Paci, P., and Doorslaer, E. van (1991). On the measurement of inequalities in health. Social Science & Medicine *33*, 545–557.

Wagstaff, A., Yip, W., Lindelow, M., and Hsiao, W.C. (2009b). China's health system and its reform: a review of recent studies. Health Economics *18*, S7–S23.

Wang, H., Du, S., Zhai, F., and Popkin, B.M. (2006a). Trends in the distribution of body mass index among Chinese adults, aged 20|[ndash]|45 years (1989|[ndash]|2000). International Journal of Obesity *31*, 272–278.

Wang, X., and Piesse, J. (2010). Inequality and the Urban-rural Divide in China: Effects of Regressive Taxation. China and World Economy *18*, 36–55.

Wang, Y., Mi, J., Shan, X. y., Wang, Q.J., and Ge, K. y. (2006b). Is China facing an obesity epidemic and the consequences? The trends in obesity and chronic disease in China. Int J Obes *31*, 177–188.

Ward, E., Jemal, A., Cokkinides, V., Singh, G.K., Cardinez, C., Ghafoor, A., and Thun, M. (2004). Cancer Disparities by Race/Ethnicity and Socioeconomic Status. CA Cancer J Clin *54*, 78–93.

Willcox, B.J., Willcox, D.C., Todoriki, H., Fujiyoshi, A., Yano, K., He, Q., Curb, J.D., and Suzuki, M. (2007). Caloric Restriction, the Traditional Okinawan Diet, and Healthy Aging. Annals of the New York Academy of Sciences *1114*, 434–455.

Williams, A.H., and Cookson, R.A. (2006). Equity-efficiency trade-offs in health technology assessment. Int J Technol Assess Health Care *22*, 1–9.

Winkelmann, L., and Winkelmann, R. (1998). Why Are the Unemployed So Unhappy?Evidence from Panel Data. Economica *65*, 1–15.

Wooldridge, J.M. (2002a). Econometric analysis of cross section and panel data (Cambridge, Mass.: MIT Press).

Wooldridge, J.M. (2002b). Inverse probability weighted M-estimators for sample selection, attrition, and stratification. Portuguese Economic Journal *1*, 117–139.

World Bank (2005). PovCalNet.

World Health Organization (1995). Physical status, the use and interpretation of anthropometry: report of a WHO expert committee (Geneva: World Health Organization).

World Health Organization (2000). World Health Report 2000 (Geneva: World Health Organization).

World Health Organization (2002). World Health Survey. Individual Questionnaire.

World Health Organization (2007). Life Tables for WHO Member States (Geneva: World Health Organization).

World Health Organization (2012a). Healthy life expectancy (HALE).

World Health Organization (2012b). World Health Statistics 2012 (Geneva: World Health Organisation).

Wu, Y. (2006). Overweight and obesity in China. BMJ *333*, 362–363.

Yang, D.T. (1999). Urban-Biased Policies and Rising Income Inequality in China. The American Economic Review *89*, 306–310.

Yip, W. (2010). Disparities in health care and health status: the rural-urban gap and beyond. In One Country, Two Societies: Rural-urban Inequality in Contemporary China, M.K. Whyte, ed. (Harvard University Press), pp. 147–165.

Yu, Y. (2008). Comment on "Reform, Growth, and Inequality in China."Asian Economic Policy Review *3*, 161–162.

Zhai, F., Wang, H., Wang, Z., Popkin, B.M., and Chen, C. (2008). Closing the energy gap to prevent weight gain in China. Obesity Reviews *9*, 107–112.

Zhang, Q., and Wang, Y. (2004). Trends in the association between obesity and socioeconomic status in U.S. adults: 1971 to 2000. Obesity Research *12*, 1622–1632.

Zhang, X., and Kanbur, R. (2005). Spatial inequality in education and health care in China. China Economic Review *16*, 189–204.

Zhong, H. (2011). The impact of population aging on income inequality in developing countries: Evidence from rural China. China Economic Review *22*, 98–107.

SUMMARY

Health differences along socio-economic lines are of all times and places. Most of the evidence points to an advantage of the rich, educated, working population. Not only do they live longer, they value their health higher and also appear healthier, measured by objective standards, such as the prevalence of (chronic) diseases. This inequality extends from rich Western countries to developing countries. It is a widely studied topic in health economics, but not of the highest priority within the health care debate in the Netherlands and other countries. International organizations such as the World Health Organization of the United Nations, the Organization for Economic Cooperation and Development (OECD) and others, put much energy in researching socio-economic health differences in their member states and consider it an important outcome of national health policy.

Despite this long-term and intensive interest in the measurement of socio-economic health inequalities, the collection of accurate and comparable information still is a daunting task. This thesis presents several challenges that researchers face. It also provides new empirical evidence on socio-economic inequalities in different areas. The results confirm the generally better health for the rich and highly educated. Moreover, the results show that with the right techniques and data many of the measurement problems that researchers face can be accounted for.

Chapter 2 describes how the measurement of socio-economic health differences may be affected by the loss of respondents in health surveys. This effect is studied in the Netherlands using a set of databases, which among other things contain information about health, income, household type and mortality. Using these data it first demonstrated how health varies between rich and poor people at different ages. A striking result is that the health differences vary strongly with age. Chapter 2 seeks to explain this variation in health inequalities by age. This is done by describing how such differences may be the result of unequal mortality for rich and poor people.

The analysis shows that unhealthy rich people are less likely to die than unhealthy poor people. As a result, in each following year there are relatively less unhealthy poor compared to unhealthy rich people. This could explain the differences described previously. This inequality in mortality is corrected for by weighing people on the basis of their survival probabilities. Respondents who resemble people who often are deceased got a high weight, while people who belong to a group with mostly survivors were given a lower weight. Although there is a difference in the survival of unhealthy rich and poor in the Netherlands, these differences are not so large that it explains the variation in health inequality between the two groups. The same reasoning and analysis can also be performed for the people who move from their own home to a nursing home or other institution. Although these people do not die, they will also not participate in the surveys

used for the measurement of health inequalities. The effects of this type of dropout out of the population are even smaller than those of mortality.

Chapter 3 deals with a different type of longitudinal study on socio-economic health differences. It focuses on how differences in health later in life are related to differences in socio-economic status during childhood. The study is situated in China and health is evaluated on the basis of the body-mass index (BMI), which is a measure of a person's body weight related to his or her height. China has in recent decades developed very quickly. The attention is usually placed on the economic developments that the country has gone through and is still going through. However, big changes have also taken place in Chinese society and the way of life. Two important examples are the how much Chinese move around nowadays and how their diet has changed. Chinese are less physically active and have taken up a less healthy diet. As a result, the weight of many Chinese is now so high that it is a significant health risk. Because obesity is now a national health problem, there is a chance that differences in activity levels and diet lead to differences in health. People with better socio-economic background might be better able to adjust their lifestyles in such a way that their health is not compromised.

This relationship could be measured directly, but a frequently studied phenomenon then presents itself. It is often not possible to know whether the health of people is determined by their socio-economic background, or that poor health prevents them from obtaining a higher socio-economic status. A third explanation might be that both outcomes do not affect each other at all, but a third, possibly unobserved, variable has an effect on both socio-economic status and health. Therefore, Chapter 3 is set-up in such a way that it is unlikely that the observed relation is a consequence of the effects of health on socio-economic status. This causal direction is unlikely, because the socio-economic status of a child is largely a result of the socio-economic status of his parents and probably not his health later in life.

Therefore, Chapter 3 measures socio-economic status as the number of years that the mother of the respondent went to school. Additionally, the chapter formulates some hypotheses on how maternal education may influence the BMI of the respondent later in life. Important explanatory mechanisms are the health of the respondent at a younger age, his or her socio-economic status later in life and the diet and activity level of the respondent. The results of these models show that there is a relationship between socio-economic status during childhood and adulthood BMI and overweight. The relationship is shown to depend on the age of the respondent. At younger ages, the children of educated mothers had a lower body weight than respondents of mother with little or no schooling. Later in life this inequality is reversed. The models that also take into account the socio-economic status of the respondent show a similar relationship. One of the main findings is that if this variable is also taken into the effects of maternal education no longer seems to be important.

Chapter 4 also studies socio-economic health inequalities in China. The link between the development of the income distribution and health is the central theme of this chapter. This relationship is important in China, because in recent decades the income distribution has changed more and faster than in other countries and in other time periods. The first step in Chapter 4 was to measure the relationship between income and health. This relationship is measured for a cohort of respondents who were followed between 1991 and 2006.

The estimated non-linear relationship between income and health was then used to measure income-related health inequalities in each available year between 1991 and 2006. The changes in these years compared to the first year (1991) were then split into four different causes: (i) increases in average income, (ii) changes in the distribution of income between individuals, (iii) shift of income ranks of healthy people compared to unhealthy people and (iv) changes in the health levels of the rich and poor over time.

The first important result of Chapter 4 is that income-related health inequalities have increased sharply between 1991 and 2006. The explanation for this increase must be sought primarily in the unequal income development of healthy and unhealthy people. It appears that this explains 70-90% of the total increase in income-related health disparities. The growth of average income and changes in income distribution itself also contribute to the increase, while the changes in health levels of rich and poor led to smaller income-related health inequalities. Further examination of the results suggests that mainly the incomes of older women, a relatively unhealthy group in society, develop less quickly than those of other, more healthy, demographic groups. Their residual income is probably the result of an inadequate pension provision for most Chinese women.

Chapter 5 compares differences in average health levels between countries. An important problem in the measurement of these differences is that these measurements are based on self-rated health. This means that a direct comparison of the reported health levels may be misleading, because not everybody defines being health in the same way. These differences in interpretation may exist between different nationalities, young and old, male and female, etc. This could have consequences for the measurement of health inequalities between rich and poor countries. Chapter 5 therefore uses existing models that correct for this subjectivity in self-rated health. These models are relatively new and section 5 introduces a number of modifications. These show that within the present approach the inequality between rich and poor countries is underestimated. The adjustments in the methodology may lead to an increase of more than 10% of the difference in healthy life years between the richest and poorest countries. Moreover, the inequality between people within a country may also be underestimated.

The measurement of health inequalities is not an end in itself, but should be aimed at supporting inequality policies. Chapter 6 discusses some methods that provide insight into the impact of reimbursement decisions on the distribution of health. As the previ-

ous chapters have shown, the measurement of health inequalities is complex. To support policy the results should at least be easy to interpret and the method used must be clear. Therefore, the first method presented in Chapter 6, just shows how the difference in average health between groups changes when some treatments are and others are not covered by health insurance. These calculations are simple and clear, but do not consider how certain health differences are valued by policymakers. Therefore, the second method considers by whom the health gains associated with a new treatment are achieved. Gains in different groups of individuals are then weight differently, based on previously elicited preferences of policymakers. These results are still easy to interpret, as they are presented in a way similar to the commonly used cost-utility ratio. However, the calculation is less transparent and other policy goals are excluded. Therefore, the third method also takes into account other criteria that are important in policy making (e.g. how often people get the disease, who get the disease, the necessity of treatment, etc.). Based on the results of discrete choice experiments outcomes of different reimbursement decisions can be compared. The three methods can lead to different conclusions and their different characteristics will determine what policy makers want to use to support their policies.

SAMENVATTING

Gezondheidsverschillen langs sociaal economische lijnen zijn van alle tijden en plaatsen. De meeste studies wijzen op een bevoordeelde positie van de rijke, goed opgeleide, werkende bevolking. Niet alleen leven zij langer, ze waarderen hun gezondheid ook hoger en blijken ook gezonder, gemeten naar objectieve maatstaven zoals de prevalentie van (chronische) ziekten. Deze ongelijkheid strekt zich uit van rijke Westerse landen tot ontwikkelingslanden. Het is een vaak bestudeerd onderwerp binnen de gezondheidseconomie, maar heeft niet de hoogste prioriteit binnen het zorgdebat in Nederland en andere landen. Internationale organisaties, zoals de Wereldgezondheidsorganisatie van de Verenigde Naties, de Organisatie voor Economische Samenwerking en Ontwikkeling (OESO) en andere, steken veel energie in het onderzoeken van sociaaleconomische gezondheidsverschillen in hun lidstaten en beschouwen de verdeling van gezondheid als een belangrijke uitkomst van nationaal zorgbeleid.

Ondanks deze langlopende en intensieve interesse in de meting van sociaaleconomische gezondheidsverschillen, blijft het vergaren van accurate vergelijkbare informatie een lastige taak. Dit proefschrift presenteert enkele uitdagingen waarmee onderzoekers geconfronteerd worden. Daarnaast levert het nieuw empirisch bewijs voor sociaaleconomisch ongelijkheid in verschillende gebieden. De resultaten bevestigen veelal de betere gezondheid voor rijken en hoogopgeleiden. Bovendien tonen de resultaten aan dat – mits de juiste technieken en data gehanteerd worden - veel van de meetproblemen waarmee onderzoekers geconfronteerd worden oplosbaar zijn.

Hoofdstuk 2 beschrijft hoe de meting van sociaaleconomische gezondheidsverschillen beïnvloed kan worden door de uitval van respondenten in gezondheidssurveys. Dit effect wordt bekeken in Nederland aan de hand van een set databases die onder andere informatie bevatten over gezondheid, inkomen, type huishouden en sterfte. Eerst wordt aan de hand van deze data uitgelegd hoe gezondheid verschilt tussen rijke en arme mensen op verschillende leeftijden. Hierbij valt op dat de gezondheidsverschillen sterk variëren met leeftijd. Hoofdstuk 2 tracht deze variatie in gezondheidsongelijkheid naar leeftijd te verklaren. Dit gebeurt door te beschrijven hoe zulke verschillen het gevolg kunnen zijn van ongelijke sterfte voor rijke en arme mensen.

Uit de analyses blijkt dat ongezonde rijke mensen minder vaak sterven dan ongezonde arme mensen. Hierdoor zijn er in elk volgende jaar relatief minder ongezonde arme mensen vergeleken met ongezonde rijke mensen. Dit zou de eerder beschreven verschillen kunnen verklaren. Deze ongelijkheid in sterfte wordt gecorrigeerd door mensen te wegen aan de hand van hun sterftekansen. Respondenten die lijken op mensen die veelal zijn overleden kregen een hoog gewicht, terwijl mensen die behoren tot een groep met veelal overlevenden een lager gewicht kregen. Alhoewel er verschil zit in de overlevingskansen van ongezonde rijken en armen in Nederland, zijn deze verschillen niet dermate groot

dat zij de variatie in gezondheidsongelijkheid tussen beide groepen kunnen verklaren. Een zelfde redenering en analyse kan ook worden uitgevoerd voor het verhuizen vanuit een eigen woning naar een verpleeghuis of andere institutie. Alhoewel deze categorie mensen niet sterft, zullen ook zij niet zichtbaar zijn in de gezondheidsdata bij het meten van gezondheidsongelijkheden. De effecten van dit type uitval uit de bevolking zijn nog kleiner dan die van sterfte.

Hoofdstuk 3 behandelt een ander type longitudinale studie naar sociaaleconomische gezondheidsverschillen. Hierin staat centraal hoe verschillen in gezondheid op latere leeftijd gerelateerd zijn aan verschillen in sociaaleconomische status tijdens de jeugd. De studie is gesitueerd in China en meet gezondheid aan de hand van de body-mass index (BMI), welke een maat is om iemands lichaamsgewicht te relateren aan zijn of haar lengte. China heeft zich in de afgelopen decennia heel snel ontwikkeld. De aandacht wordt hierbij meestal gelegd op de economische ontwikkelingen die het land heeft doorgemaakt en nog steeds doormaakt. Grote veranderingen hebben echter ook plaats gevonden in de Chinese maatschappij en manier van leven. Twee belangrijk voorbeelden hiervan zijn de mate waarin Chinezen tegenwoordig bewegen en hoe hun voedingspatroon is veranderd. Chinezen zijn minder gaan bewegen en zijn ongezonder gaan eten. Hierdoor is het gewicht van veel Chinezen nu dermate hoog dat het een belangrijk risico vormt voor hun gezondheid. Doordat overgewicht nu een nationaal gezondheidsprobleem aan het worden is, bestaat er een risico dat verschillen in beweging- en voedselpatronen leiden tot verschillen in gezondheid. Mensen met een betere sociaaleconomische achtergrond zouden wellicht beter in staat kunnen zijn hun beweeg - en voedingspatroon zo in te richten dat hun gezondheid niet in gevaar komt.

Deze relatie zou direct gemeten kunnen worden, maar een veelvuldig onderzocht probleem dient zich dan aan. Het is vaak niet mogelijk om te weten of de gezondheid van mensen beïnvloed wordt door hun sociaaleconomische achtergrond of dat het juist slechte gezondheid is, die bepaalt of mensen een lagere sociaaleconomische status hebben weten te verwerven. Een derde verklaring zou zelfs ook nog kunnen zijn dat beide uitkomsten elkaar überhaupt niet beïnvloeden, maar dat een derde (niet geobserveerde) variabele een invloed heeft op beide. Daarom is de studieopzet in hoofdstuk 3 zo, dat het onwaarschijnlijk is dat de gevonden relatie een gevolg is van effecten van gezondheid op sociaaleconomische status. De sociaaleconomische status van een kind is namelijk grotendeels een gevolg van de sociaaleconomische status van zijn ouders en waarschijnlijk niet van zijn gezondheid op latere leeftijd.

Daarom is sociaaleconomische status in hoofdstuk 3 gemeten aan de hand van het aantal jaren dat de moeder van de respondent naar school is geweest; een aantal hypotheses wordt geformuleerd hoe opleiding van de moeder de BMI van de respondent kan beïnvloeden. Belangrijke verklarende mechanismen daarbij zijn de gezondheid van de respondent op jongere leeftijd, zijn of haar eigen sociaaleconomische status op latere

leeftijd en de mate waarin de respondent beweegt en gezond eet. De uitkomsten van deze modellen tonen aan dat er een relatie bestaat tussen sociaaleconomische status tijdens de jeugd en volwassen BMI en overgewicht. De relatie blijkt wel af te hangen van de leeftijd van de respondent. Op jongere leeftijd hebben de kinderen van hoger opgeleide moeders een lager lichaamsgewicht dan respondenten van moeder met geen of weinig opleiding. Op latere leeftijd blijkt de ongelijkheid echter omgedraaid te zijn. De modellen die ook rekening houden met de sociaaleconomische status van de respondent zelf tonen een vergelijkbare relatie. Eén van de belangrijkste bevindingen daarbij is dat als deze variabele ook mee wordt genomen de effecten van de opleiding van de moeders niet meer belangrijk lijken te zijn.

Hoofdstuk 4 behandelt eveneens sociaaleconomische gezondheidsverschillen in China. De link tussen de ontwikkeling van de inkomens- en de gezondheidsverdeling staat centraal. Deze relatie is juist in China van belang, omdat de inkomensverdeling daar in de afgelopen decennia meer en sneller is veranderd dan in andere landen en in andere perioden. De eerste stap in hoofdstuk 4 was het meten van de relatie tussen inkomen en gezondheid. Deze relatie is gemeten voor een cohort van respondenten die gevolgd werden tussen 1991 en 2006.

De geschatte niet lineaire relatie tussen inkomen en gezondheid werd vervolgens gebruikt om inkomensgerelateerde gezondheidsverschillen te meten in elk beschikbaar studiejaar tussen 1991 en 2006. De veranderingen in deze jaren t.o.v. het eerste jaar (1991) werden vervolgens opgesplitst in vier verschillende oorzaken: (i) stijgingen van het gemiddelde inkomen, (ii) veranderingen in de verdeling van inkomen tussen personen, (iii) verschuiving van de inkomensrangen van gezonde mensen t.o.v. ongezonde mensen en (iv) veranderingen in de gezondheid van hoge en lage inkomens over tijd.

Het eerste belangrijke resultaat van hoofdstuk 4 is dat inkomensgerelateerde gezondheidsverschillen sterk zijn toegenomen tussen 1991 en 2006. De verklaring voor deze stijging moet voornamelijk worden gezocht in de ongelijke inkomensontwikkeling van gezonde en ongezonde mensen. Het blijkt dat dit 70 tot 90 % van de totale stijging in inkomensgerelateerde gezondheidsverschillen verklaart. De groei van het gemiddelde inkomen en veranderingen in de inkomensverdeling zelf dragen ook deels bij aan de stijging, terwijl de verandering in gezondheid van rijk en arm de inkomen gerelateerde gezondheidsverschillen juist verkleint. Verdere bestudering van de resultaten suggereert dat het vooral de inkomens van oudere vrouwen, een relatief ongezonde groep in de samenleving, zijn die achterblijven bij die van andere (gezondere) groepen. Hun achterblijvend inkomen is waarschijnlijk het resultaat van een inadequate pensioenvoorziening voor de meeste Chinese vrouwen.

Hoofdstuk 5 vergelijkt hoe de gemiddelde gezondheid verschilt tussen landen. Een belangrijk probleem bij het meten van deze verschillen is dat deze metingen gebaseerd zijn op zelf gewaardeerde gezondheid. Dat betekent dat een directe vergelijking van de

gerapporteerde gezondheid misleidend kan zijn, omdat niet iedereen hetzelfde verstaat onder gezond en ongezond zijn. Deze verschillen in interpretatie kunnen bestaan tussen verschillende nationaliteiten, oud en jong, man en vrouw, enz. Dit kan dus gevolgen hebben voor het meten van gezondheidsongelijkheid tussen rijke en arme landen. Hoofdstuk 5 gebruikt daarom bestaande modellen die corrigeren voor deze subjectiviteit in zelf gewaardeerde gezondheid. Deze modellen zijn nog relatief nieuw en hoofdstuk 5 introduceert daarom een aantal aanpassingen. Hieruit blijkt dat de huidige aanpak de ongelijkheid tussen rijke en arme landen onderschat. De aanpassingen in de methodologie kunnen leiden tot een toename van meer dan 10% in het verschil in gezonde levensjaren tussen de rijkste en armste landen. Bovendien kan ook de ongelijkheid tussen mensen binnen een land onderschat worden.

Het meten van gezondheidsongelijkheid is geen doel op zich, maar zou er op gericht moeten zijn om ongelijkheidsbeleid te ondersteunen. Hoofdstuk 6 bespreekt een aantal methoden die, bij besluiten rondom vergoeding van (nieuwe) zorgbehandelingen, inzicht verschaffen in de impact op de verdeling van gezondheid. Zoals de eerdere hoofdstukken aantoonden, kan het meten van gezondheidsongelijkheid ingewikkeld zijn. Om beleid goed te kunnen ondersteunen, moeten de resultaten op zijn minst eenvoudig te interpreteren zijn en de gebruikte methoden inzichtelijk. Daarom toont de eerste in hoofdstuk 6 gepresenteerde methode, slechts hoe het verschil in gemiddelde gezondheid tussen groepen verandert als sommige behandelingen wel en niet vergoed worden. Deze berekeningen zijn eenvoudig en inzichtelijk, maar houden geen rekening met hoe bepaalde gezondheidsverschillen worden gewaardeerd door beleidsbepalers. Daarom weegt de tweede methode de gezondheidswinsten, die behaald worden met de nieuwe behandeling, aan de hand van wie zijn gezondheid verbetert. De preferenties die deze weging bepalen, zijn daarbij in een eerder stadium al gemeten. Deze resultaten zijn nog altijd eenvoudig te interpreteren, aangezien deze vergelijkbaar zijn met de veel gebruikte kosten-utiliteitsratio. De berekening is echter minder inzichtelijk en andere beleidsdoelen blijven buiten beschouwing. Daarom wordt in de derde methode ook rekening gehouden met andere criteria die meegenomen worden door beleidsmakers (bijv. hoe vaak mensen de ziekte krijgen, wie de ziekte krijgen, de noodzakelijkheid van behandeling, enz.). Op basis van keuzenexperimenten is de voorkeur voor effecten op de verschillende criteria gemeten en kunnen verschillende behandelingen met elkaar vergeleken worden. De drie methoden kunnen leiden tot verschillende conclusies en hun verschillende eigenschappen zullen bepalen wat beleidsmakers willen gebruiken ter ondersteuning van hun beleid.

DANKWOORD

Tijdens een promotietraject wordt je geacht jezelf te ontwikkelen tot een zelfstandige onderzoeker. Dat punt zou ik nu dus bereikt moeten hebben. In de afgelopen jaren heb ik echter zo veel goede onderzoekers leren kennen, dat ik dit nog altijd als iets heel relatiefs zie. Bovendien zou ik zonder de hulp van een aantal van hen, maar ook van vrienden en familie, nooit zijn gekomen waar ik nu ben.

Allereest moet ik Tom, zonder wie ik de eindstreep nooit had kunnen halen, natuurlijk bedanken. Jouw kennis, toewijding en consciëntieuze manier van werken hebben mij vanaf de eerste dag versteld doen staan. Dat ik door een samenloop van omstandigheden uiteindelijk jouw eerste PhD student mocht zijn, heeft voor mij heel positief uitgepakt. Ik en mijn compleet andere manier van werken maakten het je niet altijd makkelijk. Toch heb je dit nooit de samenwerking laten beïnvloeden.

De extra hulp van Eddy is ook onontbeerlijk geweest. De oorspronkelijke afspraak om op afstand promotor te zijn, moesten we al snel aanpassen. Na een paar weken moest ik namelijk al op zoek naar een nieuwe begeleider. De flexibiliteit en oplossingsgerichtheid die je toen hebt laten zien, zijn ook later in het traject nog vaak van belang geweest.

Teresa, without you it would have been impossible to finish Chapter 5. It may have been one of the most ungrateful papers to have been a co-author of. The dedication and helpfulness you displayed in the last months, even into your maternity leave, to help me finish this thesis are remarkable.

Hoe fijn het ook is om het proefschrift nu af te ronden, toch kijk ik nog altijd met veel plezier terug op het begin. Daarom wil ik Louis en Xander bedanken die mij de basis van onderzoeken hebben geleerd en bij BMG hebben geïntroduceerd.

Dennis, heel erg bedankt voor je hulp met het ontwerp van dit boekje. Leuk om te zien dat ook hierbij de kleine details het verschil kunnen maken.

Bij GE-iMTA was er gelukkig altijd ruimte voor meer dan alleen werk. Hierdoor kijk ik terug op een hele mooie periode en mag ik veel van mijn oud-collega's nog altijd vrienden noemen. Ik wil dan ook iedereen bedanken die mee heeft gedaan aan de reisjes voor of na een congres, de Roparun, één van de vele borrels of gewoon een kopje koffie kwam drinken na de lunch.

Toch moet ik twee mensen specifiek bedanken: mijn paranimfen Caroline en Liesbet. Caroline, we waren bijna vier jaar lang kamergenoten. Ik ben dankbaar dat je al snel mijn gemompel tijdens het werken wist te negeren en altijd klaar stond om te luisteren naar mijn verhalen, frustraties en andere onzin die even belangrijker leken dan jouw of mijn werk. Liesbet, ook na een paar maanden blijft het raar elkaar niet meer elke dag te zien. Ik mis niet alleen onze cc-tjes, maar ook je vrolijkheid en nuchtere kijk op de dingen van alledag. Ik ben heel blij dat we 11 jaar als studiegenoten en collega's afsluiten samen voor de commissie.

Ten slotte zijn de waardering, begrip en geduld van mijn familie van onschatbare waarde geweest. Joost, je bent altijd een voorbeeld geweest om het beste uit mezelf te halen. Ook bij deze promotie heeft dat zeker geholpen. Lieve Ernest en Marja, het is onmogelijk om jullie voldoende te bedanken, zeker in deze paar regeltjes hier. Bij alles wat ik altijd gedaan heb, is jullie steun onvoorwaardelijk geweest. Het nu afgesloten traject is daar geen uitzondering op geweest. Jullie interesse in waar ik nou al die tijd mee bezig was, hield ik altijd een beetje af. Toch heeft jullie belangstelling altijd zeer motiverend gewerkt en bijgedragen aan een succesvolle afronding.

CURRICULUM VITAE

Steef Baeten was born in Nieuwegein on June 7th 1982. After finishing high school in 2001 he started his studies at the Erasmus University Rotterdam. After obtaining his BSc in Health Policy and Management, he did a two year research master in Clinical Epidemiology, with a specialization in Health Services Research, at the Neterherlands Institute for Health Sciences (NIHES). He then continued his studies at Erasmus University at the institute for Health Policy and Management (iBMG) by enrolling in the master Health Economics, Policy and Law (HEPL) with a specialization in Health Economics. He graduated in 2008.

During his period as a student of the HEPL program Steef started working as junior researcher at the institute for Health Technology Assessment (iMTA). He worked for one year on the Komen breast cancer project, before becoming a PhD student at the iBMG in 2008. During his PhD he supervised computer workgroups in the HEPL course Economics of Health and Health Care. He also worked as an instructor in the courses Economics of Ageing and the Research Project in the (international) bachelor program of the Erasmus School of Economics.

As of December 2012 Steef works as a junior research consultant at Pharmerit.

PhD PORTFOLIO

Courses

Erasmus Summer Programme - Netherlands institute for health sciences
- Principles of research in medicine
- Introduction to public health
- Methods of health services research
- Methods of public health research
- Health economics
- Economic evaluation
- Topics in meta-analysis
- Decision making in medicine
- Pharmaco-epidemiology
- Prevention research
- Topics in evidence based medicine

Other courses - Netherlands institute for health sciences
- Study design
- Classical methods for data-analysis
- Modern statistical methods
- Analysis of population of health
- Analysis of determinants
- International comparison of health care systems
- Economic evaluation
- Research on quality of care
- Evidence-based policy making
- Qualitative research

Other courses
- Academic writing

Teaching

Economics of Health and Health Care - *Master Health Economics Policy and Law*

Economics of Ageing - *International Bachelor Economics and Business Economics*

Research Project - *International Bachelor Economics and Business Economics/Bachelor Economie*

MSc Thesis co-reader - *Master Health Economics Policy and Law*

Conferences

European Conference on Health Economics. 2012. Zürich, Switzerland. (Presenter)

Health, Development and Inequality. 2012. Darmstadt, Germany. (Work discussed)

8th World Congress on Health Economics. 2011. Toronto, Canada. (Presenter)

3rd Low Lands Health Economists' Study Group. 2011. Soesterberg, the Netherlands. (Attendant)

2nd The Low Lands Health Economists' Study Group. 2010. Egmond aan Zee, the Netherlands. (Work discussed)

7th World Congress on Health Economics. 2009. Peking, China. (Presenter)

Other publications

Bouvy, J.C., Fransen, P.S.S., Baeten, S.A., Koopmanschap, M.A., Niessen, L.W., & Dippel, D.W.J. (2012) Cost-effectiveness of two endovascular treatment strategies vs intravenous thrombolysis. In: *Acta Neurologica Scandinavica*.

Dirks, M., Baeten, S.A., Dippel D.W.J., Exel, N.J.A. van, Wijngaarden, J. van, Huijsman, R., Koudstaal, P.J., Niessen, L.W. (2012) Real-life costs and effects of an implementation program to increase thrombolysis in Stroke. In: *Neurology* E-pub 2012 July 11.

Ginsberg, G.M., Lauer, J.A., Zelle, S., Baeten, S., Baltussen, R., 2012. Cost effectiveness of strategies to combat breast, cervical, and colorectal cancer in sub-Saharan Africa and South East Asia: mathematical modelling study. In: *BMJ* 344, e614–e614.

Baeten, S.A. Baltussen, R.M.P.M., Uyl - de Groot, C.A., Bridges, J.F.P., Niessen, L.W. (2011). Reducing Disparities in Breast Cancer Survival - The Effect of Large-Scale Screening of the Uninsured In: *The Breast Journal*, (Eds.), 17(5), 2 pp., 548-549

Baeten, S.A., Baltussen, R.M.P.M., Uyl - de Groot, C.A., Bridges, J., Niessen, L.W. (2010). Incorporating Equity-Efficiency Interactions in Cost-Effectiveness Analysis-Three Approaches Applied to Breast Cancer Control In: *Value in Health*, (Eds.), 13(5), 7 pp., 573-579

Baeten, S.A., Exel, N.J.A. van, Dirks, M., Koopmanschap, M.A., Dippel, D.W.J., Niessen, L.W. (2010). Lifetime health effects and medical costs of integrated stroke services - a non-randomized controlled cluster-trial based life table approach In: *Cost Effectiveness and Resource Allocation*, (Eds.), 8(21), 10 pp., 1-10