O'Donnell, Owen; Doorslaer, Eddy Van; Ourti, Tom Van

Working Paper

Health and Inequality

Tinbergen Institute Discussion Paper, No. 13-170/V

Provided in Cooperation with:
Tinbergen Institute, Amsterdam and Rotterdam

Suggested Citation: O'Donnell, Owen; Doorslaer, Eddy Van; Ourti, Tom Van (2013) : Health and Inequality, Tinbergen Institute Discussion Paper, No. 13-170/V, Tinbergen Institute, Amsterdam and Rotterdam

This Version is available at:
http://hdl.handle.net/10419/87382

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
Health and Inequality

Owen O’Donnell\textsuperscript{a,b,d}
Eddy Van Doorslaer\textsuperscript{a,c,d}
Tom Van Ourti\textsuperscript{a,d}

\textsuperscript{a} Erasmus School of Economics, Erasmus University Rotterdam, The Netherlands;
\textsuperscript{b} University of Macedonia, Greece;
\textsuperscript{c} Institute for Health Policy and Management, Erasmus University Rotterdam, The Netherlands;
\textsuperscript{d} Tinbergen Institute, The Netherlands.
Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

More TI discussion papers can be downloaded at http://www.tinbergen.nl

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 1600

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Duisenberg school of finance is a collaboration of the Dutch financial sector and universities, with the ambition to support innovative research and offer top quality academic education in core areas of finance.

DSF research papers can be downloaded at: http://www.dsf.nl/

Duisenberg school of finance
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 8579
Abstract

We examine the relationship between income and health with the purpose of establishing the extent to which the distribution of health in a population contributes to income inequality and is itself a product of that inequality. The evidence supports a significant and substantial impact of ill-health on income mainly operating through employment, although it is difficult to gauge the magnitude of the contribution this makes to income inequality. Variation in exposure to health risks early in life is a potentially important mechanism through which health may generate, and possibly sustain, economic inequality. If material advantage can be exercised within the domain of health, then economic inequality will generate health inequality. In high income countries, the evidence that income (wealth) does have a causal impact on health in adulthood is weak. But this may simply reflect the difficulty of identifying a relationship that, should it exist, is likely to emerge over the lifetime as poor material living conditions slowly take their toll on health. There is little credible evidence to support the claim that the economic inequality in society threatens the health of all its members, or that relative income is a determinant of health.

JEL code: D31, I14, J3

Keywords: income, wealth, health, inequality

Acknowledgements: We thank the editors, Tony Atkinson and François Bourguignon, for their encouragement and detailed comments on a draft of this chapter. Thanks are also due to Hans van Kippersluis and participants of the NH Handbook of Income Distribution conference, particularly Andrew Clark, for suggestions. Hale Koc and Hao Zhang provided excellent research assistance. We are grateful to various organisations for collecting and providing data (see Appendix). We acknowledge support from the National Institute on Aging, under grant R01AG037398.
1. Introduction

The financially better-off tend also to be in better health. This holds between and within countries, both developed and developing, and has been evident for a considerable period of time (Hibbs 1915, Woodbury 1924, Van Doorslaer, Wagstaff et al. 1997). There is an income gradient in mortality, as well as a variety of measures of morbidity and disability. The relationship is not confined to a health gap between the poor and the rest; health continues to rise with income among the non-poor.

The strength, ubiquity and persistence of the positive relationship between income and health make it of considerable interest from the perspectives of the distributions of income, health and well-being. Understanding the mechanisms that generate the income-health nexus can help account for inequality, as well as identify inequity, in each of those distributions. This chapter examines the strength and nature of the relationships between income and health with the purpose of establishing the extent to which the distribution of health in a population contributes to economic inequality and is itself a product of that inequality.

The distribution of health is potentially both a cause and a consequence of the distribution of income. Differences in health can generate differences in income, most obviously by restricting earnings capacity. But health inequality may itself reflect economic inequality if health enhancing goods, such as medical care and nutritious food, are allocated by price. The potentially bidirectional relationship between health and income is relevant both to the positive explanation of the distribution of income and to its normative evaluation. A full understanding of how income differences across individuals are generated requires identification of the extent to which health constrains income. This positive exercise feeds into the normative one of evaluating the distribution of income since the inequity of income inequality surely depends on its causes. Ethical judgment of the income distribution is also contingent on its consequences. If money can buy health, then there may be greater aversion
to inequality in the distribution of income than there would be if the rich were merely able to afford smarter clothes and faster cars.

The relationship between income and health is not only of interest to those concerned with the distribution of income. From the public health perspective, attention is drawn to observed increases in health with income, as opposed to the corresponding decrease in income with ill-health. Public health scientists tend to interpret the income gradient in health as a symptom of inequity in the distribution of health (Commission on the Social Determinants of Health 2008), while economists are inclined to view the gradient as reflecting the operation of the labour market in which the sick and disabled are constrained in capacity to generate earnings (Smith 1999, Deaton 2002, Smith 2004). Resolution of this debate is obviously crucial to formation of the appropriate policy response to the gradient. If it mainly reflects the impact of ill-health on income, then the proposal that is sometimes made to use income redistribution as an instrument of health policy (Navarro 2001, Commission on the Social Determinants of Health 2008) would be entirely inappropriate (Deaton 2002).

Inclusion of this chapter in this Handbook is partly motivated by insights into the explanation and evaluation of the distribution of income that can be gained by study of the income-health relationship, but it also reflects a trend away from the more narrow focus on differences in income to the more encompassing analysis of inequality in well-being. Health and income are typically cited as the most important determinants of well-being and are the most common arguments of multi-dimensional measures of inequality (see chapters 3 and 4 of this volume). For given degrees of inequality in the marginal distributions of income and health, most would consider that inequality in well-being is greater when the poor also tend to be in worse health. Understanding the nature of the relationship between income and health is central to determining the degree of inequality in well-being.
Health to income

There are multiple mechanisms through which health may impact on the income distribution, with the labour market obviously being an important channel. Differences in productivity deriving from variation in physical and mental capacities related to illness and disability are potentially important determinants of earnings. Differences in the nature of work and infrastructures may mean that physical disability represents a greater constraint on earnings in low income settings, while mental health problems are relatively more important in developed countries. Discrimination may further widen any disparity in earnings between the disabled and able-bodied. Institutional constraints on wage flexibility may result in unemployment of less healthy individuals of lower productivity or facing discrimination. On the supply side, ill-health may shift preferences away from work and this may be reinforced by reduced financial incentives arising from a lower offer wage and entitlement to disability insurance. The latter will cushion the earnings loss arising from disability and so compress the income distribution but this will be offset if the financial incentives induce withdrawal from employment at a given degree of disability, which may strengthen the earnings-health, if not the income-health, relationship in high- relative to low-income countries. Beyond its effect on the distribution of personal income, health may impact on the distribution of household income through the formation and maintenance of marriage partnerships, and spousal earnings given needs for informal care.

The impact of health on income may operate with a very long lag. Poor health in childhood may disrupt schooling. Exposure to health risks in utero and illness in infancy may impair cognitive functioning and reduce the efficiency of education in producing knowledge and skills. Childhood health problems may be persistent, such that less healthy young adults enter the labour market with less human capital and lower lifetime earnings prospects. Early life health conditions may impact on income not only through human capital acquisition but also
by triggering health problems in adulthood (Barker 1995) that interfere with work. If exposure to health risks in early life is related to economic circumstances, then childhood health could be partly responsible for the transmission of these circumstances across generations (Currie 2009). According to this proposition, poorer mothers with less education deliver less healthy babies and raise sicker children who acquire less human capital and suffer persistent health problems, both of which constrain earnings and increase the likelihood of parenting a child with health problems. If this theory is empirically significant, then it would place health policy at the very heart of social policy.

**Income and income inequality to health**

The distribution of income potentially has consequences for population health through two broad mechanisms. First, the health of an individual may depend on his or her (parents’) level of income. If health is a normal good, then demand for it rises with income and the relationship should be stronger in countries that rely more on the market to allocate health resources, in particular medical care. Second, some claim that the health of an individual is contingent not only on his own income but also on the economic inequality within the society in which he lives (Wilkinson 1996, Wilkinson, Pickett 2010). Aggregate data show a clear negative association between measures of population health and income inequality. One proposed mechanism is that psychosocial stress arising from the stigma attached to low relative incomes is physiologically damaging. But the negative relationship between average health and income dispersion could also arise from decreasing health returns to absolute income (Rodgers 1979, Gravelle 1998). We weigh the evidence that not only income but also income inequality has a causal impact on health and so affects the distribution of health in a population.
Scope of the chapter

The literature on the socioeconomic determinants of health is immense and comes from epidemiology, sociology, demography and psychology, as well as economics. We confine attention to the relationship between income and health, which has been the focus of the economics discipline. Our goal is to establish what is known of the relationship from empirical analyses, and we do not cover the normative literature on health inequality. The evidence we assemble is relevant to ethical judgment of the distributions of income, health and well-being but we do not discuss how such normative evaluations might be conducted. Interested readers can consult the excellent discussion of some of the normative issues by Marc Fleurbaey and Erik Schokkaert (2011), as well as chapters 3 and 5 of this volume.

Related to this, we do not cover the burgeoning literature on the measurement of income-related health inequality (Van Doorslaer, Van Ourti 2011, Erreygers, Van Ourti 2011). Inequality in both income and health could also be analyzed using measures of multidimensional inequality, which are discussed in chapter 4 of this volume.

Population health is a standard covariate in empirical growth models and its contribution to growth has been the focus of a substantial literature aiming to estimate the economic returns to health investments (Commission on Macroeconomics and Health 2001, Barro 2003, Barro 2013). We do not cover this literature on the relationship between average income and health since it says nothing about the distribution of each variable across individuals. We do cover evidence on the impact of individual health on income, and of income on health, from low income, as well as high income, countries. But the balance is tilted toward a focus on the latter. Covering the very large literature on the impact of health (and nutritional status) on earnings in low income settings (Strauss, Thomas 1998) comprehensively would be too unwieldy. We refer to this literature mainly to establish whether the income-health
relationships observed in this setting differ from those that we consider in more detail from high income economies.

While we have referred until now to the relationship between income and health, our scope is a little broader. We also consider the relationships between wealth and health. Wealth is an economic outcome of intrinsic interest and is arguably a more appropriate indicator of the economic status of older individuals who provide much of the action in terms of variation in health. The health-wealth effect is likely to differ from the health-income effect. Health may affect income largely through labour market returns. This will feed through to the distribution of wealth but, in addition, ill-health may threaten wealth through asset depletion to pay for medical and nursing care.

**Organization of the chapter**

We begin by illustrating the strong positive relationship between health and income using data from three countries – China, the Netherlands and the United States – that differ greatly with respect to level of development, economic inequality, labour market structures and social welfare institutions. For each country, we show the contribution (in a purely statistical sense) that health differences make to income inequality, and, from the other side, the extent to which income variation accounts for health inequality. Having established the strength of the association between income and health, in subsequent sections we turn to the mechanisms potentially responsible for the relationship and the extent to which it arises from a causal effect of health on income, and vice versa. Section 3 identifies a number of routes through which health may impact on income and wealth, paying particular attention to how economic inequality may be generated by health differences. The pathways considered are wages, work, human capital, early life health risks, occupation, marriage and medical expenditures. Evidence relevant to each broad pathway and more specific mechanisms is reviewed. Section 4 looks at the relation from the other direction: income (wealth) to health. Much of the
discussion is concerned with whether income (wealth) has a causal impact on health over and above that of other socioeconomic characteristics, such as education and occupation, and after controlling for correlated determinants, such as time preferences and risk attitudes. Section 5 considers the logic and empirical support for the hypothesis that health is determined by economic inequality and by relative, as opposed to absolute, income. The final section briefly summarizes the lessons that can be drawn from the literature about the nature of the income-health relationship and discusses what these imply for the normative evaluation of the distributions of income, health and well-being.

2. Health and income: a first pass

To whet the appetite, we illustrate the strength of the relationship between health and income in the United States (US), the Netherlands (NL) and China.\(^1\) The purpose is simply to show that there is a substantial and ubiquitous relationship that deserves attention, and to assess its potential relevance to explanation of inequalities in the distributions of income and health. The three countries are chosen primarily because of their differences. One is large, rich but unequal, and does not (yet) have universal health insurance coverage. Another is small, rich and egalitarian, and provides universal health coverage and extensive social protection, including disability insurance, typical of northern continental Europe. The third is very large, much poorer (but rapidly becoming less poor) and less healthy than the other two, with increasing economic inequality, and limited health and disability insurance coverage. Differences in the wealth and economic structures of these countries, as well as their health

---

\(^1\) The US data are from the 2008 well-being module of the American Life Panel (ALP), which is nationally representative and implemented by RAND over the internet (https://mmicdata.rand.org/alp/). The Dutch data are from the 2011/12 wave of the Longitudinal Internet Studies for the Social sciences (L ISS), which is also nationally representative and has a similar protocol to the ALP (http://www.lissdata.nl/lissdata/). The Chinese data are from the 2006 wave of the Chinese Health and Nutrition Survey (CHNS), which is representative of eight provinces not including the megacities on the eastern seaboard (http://www.cpc.unc.edu/projects/china/).
and welfare institutions, might be expected to be reflected in the distributions of income and health, as well as the association between them.

Figure 18.1 illustrates the income gradient in self-assessed health (SAH) (Smith 2004) – the most common survey measure of general health that invites a respondent to select one of four (China) or five (NL & US) labels as the best description of his or her health. We focus on the percentage reporting less than *good* health, which always corresponds to the bottom two categories of SAH, by age-specific quartiles of household per capita income. In the United States, this percentage rises monotonically as income falls at all ages except among the oldest (70+). Even the poorest elderly, whose income should not depend on their current health, are more than twice as likely as their richest contemporaries to report less than *good* health. The pattern is similar in the Netherlands but for the absence of a gradient among young adults and a weaker gradient among the elderly. In both countries, the gradient increases until middle-age and narrows beyond that when retirement becomes more prevalent. This is consistent with employment being an important characteristic linking health to income. At the peak of the gradient between the ages of 50 and 59, more than 40% of the poorest Americans report their health to be less than *good* compared with less than 10% of their richest compatriots. The inequality is narrower in the Netherlands but the poor middle-aged are still around three-and-a-half times more likely to report less than *good* health than are those in the top quartile of the income distribution.

In China, the main health disparity is not between the poorest quartile and the rest, as it is in the US and the Netherlands, but, if anything, between the richest quartile and those less privileged. There is no narrowing of the gradient in old age in China. In fact, health differences are greatest in the oldest age group, which is consistent with sizable inequalities in pension entitlements and health insurance coverage among the Chinese elderly but may also
reflect the fact that the Chinese survey asks respondents to report health relative to others of the same age (see note to Figure 18.1).

Figure 18.1: Percentage reporting less than *good* health by age-specific quartiles of household per capita income in United States (US), the Netherlands (NL) and China

Notes: Authors calculations from 2006 CHNS (China), 2011 LISS (NL) and 2008 ALP (US). ALP and LISS respondents report health as being excellent, very good, good, fair (ALP)/ moderate (LISS), or poor. CHNS respondents report health *relative to others of their own age* as very good, good, fair or poor. Household income is before payment of taxes and social security contributions and after receipt of transfers. Annual income for ALP and monthly income for LISS and CHNS, all in local currencies.

Figure 18.2 shows the flip side of the relationship between health and income. In all three countries, those in (at least) *very good* health have substantially higher incomes than those in *poor* health. In the US at all ages, mean household per capita income falls as health drops from one category to the next. The health gradient in income peaks in the prime years of working life (40-49), when the mean income of those reporting *excellent* or *very good* health is around three-and-a-half times greater than that of those in poor health. Even in old age, those with the best health have almost twice the incomes of those with the worst health. The health-related income gaps are narrower in the Netherlands. Even at the ages (50-59 years) where the disparity is greatest, those in the best health do not receive twice the income of those in the worst. The relative income differences by health in China are similar in magnitude to those in the Netherlands, except in old age when the gap widens, rather than narrows.
Figure 18.2: Mean household per capita income by self-assessed health and age, United States (US), the Netherlands (NL) and China

Notes: As Figure 1.

The health-income relationship remains strong after controlling for some potential correlates of both. Table 18.1 shows estimates from least squares regressions of the logarithm of household equivalent income on SAH, gender, age, ethnicity, education, and region (US and China). Conditional on these characteristics, in the US sample the mean income of those reporting very good or excellent health is approximately 66-69% greater than that of someone reporting poor health (left panel, 1st column). This is larger than the relative income difference between those with the middle (post-high school vocational) and lowest (high school graduate or less) level of education, although it is not quite as large as the difference between university graduates and those with no more than high school education. Controlling for employment status has a very large impact on income differences by health. In the US, the mean income of those with at least very good health relative to those in poor health is reduced by half to 32-33% (right panel, 1st column). A large part of the strong relationship between income and health appears to be mediated through employment.

Multivariate analysis confirms what is suggested by Figure 18.2 – that income gaps by levels of health are narrower in the Netherlands and China than in the US. Without conditioning on employment, the Dutch reporting at least very good health have incomes approximately 31-

---

2 Definitions of household equivalent income and the covariates are provided in Appendix Table A1.
37% higher than their compatriots in poor health. This is much lower than the respective relative disparity in the US sample. But conditioning on employment has a much smaller impact than it does in the US, reducing the difference by around a quarter to 23-27%, which is only slightly less than the relative income difference of 30% between those with the highest and lowest levels of education. The more modest effect of conditioning on employment may be a reflection of the more generous disability insurance in the Netherlands, which is evident in the coefficients on disability insurance status. Conditioning on employment has little or no impact in the Chinese sample. Those reporting very good health have incomes approximately 36% higher than those in poor health. This is because employment differs less by health in China (see Figure 18.4).

Large differences in income by health do not necessarily imply that health statistically explains, let alone causally determines, a substantial part of income inequality. Whether it does depends on the degree of health variation that exists in the population, in addition to its partial correlation with income. The percentage of respondents reporting poor health is only 1.5% in the Dutch sample, rising to 3% in the US and to 7% in the Chinese (Appendix Table A1). Differences in income between those with poor and higher levels of health may therefore make only modest contributions to the explanation of income inequality.

To give an impression of the contribution that health can make to the explanation of income inequality, we apply a simple version of a Shapley value approach (Sastre, Trannoy 2002, Shorrocks 2013) to decompose the relative Gini index estimated from the three datasets. This involves writing the income of each individual as the predicted value from the regression, plus the residual, and calculating the marginal impact on the Gini of neutralising a variable by fixing its value across all individuals. This impact will vary depending on the covariates that
have previously been held constant. The contribution of a variable to inequality is given by its average marginal impact across all possible sequences of neutralising the set of all covariates.³

As would be expected, the estimated Gini indices reveal inequality to be lowest in the Netherlands (0.29), and of similar magnitude in the US (0.46) and China (0.47). The percentage contributions of the factors to income inequality are reported adjacent to the respective column of regression coefficients in Table 18.1. Without conditioning on employment status, SAH explains 6.5% of income inequality in the US sample (left panel, 2nd column). This is similar in magnitude to the contribution of race, a little less than that of age and about one third of that of education. The decomposition obviously depends on the specification of the regression model. Once employment status is added to the controls, health only explains 4% of income inequality, which is half of the contribution of age and a little more than one quarter of that of education.

Without conditioning on employment, health explains 3.6% and 3.0% of income inequality in the Netherlands and China respectively. In each country, this is substantially more than the contribution of both gender and ethnicity. The health contribution is around twenty percent higher than that of age in China, but less than the age contribution in the Netherlands. In both countries, as in the US, the contribution of variation in SAH to the explanation of income inequality is substantially less than that of education. Predictably from the estimates of the

³ Income of individual $i$ is given by $\exp(\tilde{y}_i + \sum_{k=1}^{K} \hat{\gamma}_k x_{ki} + \hat{\epsilon}_i)$ where $\hat{\gamma}_k$ are coefficients from the OLS regression of log income and $\hat{\epsilon}_i$ is the residual. Inequality in this measure is computed and compared for different combinations of the regressors ($x_{ki}$) and residuals fixed at particular values. We compute the aggregate contribution of a factor, such as SAH, that consists of several categories represented by dummy variables, and do not attempt to establish the contribution of each separate category. Use of the relative Gini and a regression model of log income ensures that the decomposition is insensitive to the values at which the regressors are fixed and to the estimate of the constant ($\hat{y}_0$).
regression models, controlling for employment status has less impact on the extent to which health explains income inequality in the Netherlands and China than it does in the US.

Without conditioning on employment, which is the most obvious route through which health influences income variation in SAH accounts for 6.5%, 3.6% and 3.0% of income inequality in the US, the Netherlands and China respectively. While these contributions appear modest, one must bear in mind that most inequality remains unexplained by an admittedly rather restricted set of factors in all three countries. SAH accounts for almost 15% of the explained income inequality in the US and the Netherlands, and just over 10% in China. In addition, SAH is only one measure of health and varies only over four or five categories. It is inevitable that such a variable will not be able to account for a substantial proportion of the variation in continuous incomes. Differential reporting of health may also result in SAH understating the correlation between income and health (Bago d’Uva, van Doorslaer et al. 2008). While it is established that SAH is an informative summary measure of health, the addition of more health measures, particularly continuous ones and those capturing disabilities, to the decomposition analysis would inevitably increase the proportion of income inequality that is explained by health variation.

From the regression and decomposition analysis presented nothing can be inferred about causality — its direction or even existence. One can just as well look at health differences that are explained by income variation. Table 18.2 presents estimates from interval regressions of transformations of SAH on household equivalent income and the same covariates used in the income regressions (Doorslaer, Jones 2003). This health measure lies between 0 (minimum health) and 1 (maximum health). In the US, the difference in predicted health between the

---

4 Thresholds separating different categories of SAH are taken from external data. For China, we use the Chinese visual analogue scale estimated from the World Health Organization Multi-Country Survey on Health and Responsiveness (Üstün, Chatterji et al. 2003), which gives thresholds 0 (minimum health), 0.50, 0.80, 0.91, and 1 (maximum health) corresponding to four categories of SAH. For the US and the Netherlands, we obtain the thresholds 0, 0.428, 0.756, 0.897, 0.947, and 1 from the Canadian Health Utility Index (Feeny, Furlong et al. 2002) that divide five categories of SAH.
richest and poorest quartile of household equivalent income is about 1.8 times greater than the
difference between the youngest and oldest age group, and between the highest and lowest
education categories (left panel, 1st column). As could be anticipated from the narrower
income gaps by health in the Netherlands observed in Table 18.1, the health differences by
income are also smaller than those in the US (Table 18.2, middle panel). Still, the health
disparity between the richest and poorest income groups is more than twice the difference
between the top and bottom education groups. In China, the health differences by income are
also substantially larger than the differences by education.5

Table 18.2 also presents results from Shapley value decompositions of inequality in
transformed SAH. There are two differences from the decomposition of income inequality to
note. First, the categorical nature of SAH, modeled by interval regression, means that
unexplained variation in health cannot be captured by the decomposition. All contributions
refer to the percentage of the explained variation that is accounted for by a factor. Second, the
absolute, rather than relative, Gini index (Yitzhaki 1983) is a more appropriate measure of
inequality in a bounded variable like transformed SAH (Erreygers 2009, Lambert, Zheng
2011).6

Income quartiles account for 45% of the explained inequality in SAH in the US (left panel, 2nd
column). This is roughly equal to the contributions of age and education combined. Further
evidence that employment is central to the association between health and income in the US is
provided by the fall in the income contribution by more than half, such that it becomes only
slightly more important than education in explaining health inequality, when employment

5 Conditioning on employment status reduces the health difference between the richest and poorest by about two-
fifths in the US. Doing so results in a smaller reduction in the income gradient in health in the Netherlands,
which remains about twice the education gradient. Conditioning on employment has little impact in the Chinese
sample.
6 Since the interval regression model has an additive specification and we use an inequality index that is invariant
to equal additions, neither the constant nor the values to which factors are set when neutralized will affect the
decomposition results.
status is added to the interval regression of SAH and so the decomposition (left panel, 3rd column). Adding proxies for health behaviour, in the form of indicators of weight and smoking, reduces the income contribution by about one third more, which is suggestive of differences in lifestyle being an important reason why health differs by income.

In the Netherlands, income variation accounts for 35% of the explained inequality in SAH, which is more than twice the contribution of education. Adding employment status reduces the contribution of income by half but it remains roughly twice that of education. Entering health behaviour into the decomposition has a more modest impact. In China, income differences account for much less of the explained variation in health; around 9% irrespective of whether employment or health behaviour are controlled for. As is also evident in Figure 18.1, health inequality in China appears to be driven mainly by age, which is perhaps surprising given that only the Chinese survey asks respondents to report their health relative to that of someone of the same age. On the other hand, a very steep decline in health with age in China would be anticipated by the lack of health insurance coverage, particularly for the elderly in rural locations, at the time of the survey, as well as from the events of recent Chinese history.

[Table 18.2 here]

Our empirical illustration demonstrates that there is a strong relationship between income and health. Income differences by health are large. Correspondingly, health disparities by income are wide. The relationship is stronger in the US than in the Netherlands and China. The dissimilarities of the latter two countries implies that the strength of the relationship is not explained by a simple factor, such as the informality of the economy, universality of health insurance cover, or the generosity of welfare, but is likely a product of many such factors. Employment status, particularly in the US among the three countries examined, is key to the
relationship between income and health. But it is unlikely to be the only mechanism. Even after controlling for employment, income differs greatly by health and so health varies substantially with income.

A single health variable (SAH) accounts for 6.5% of total income inequality and 14.6% of explained inequality in the US. These estimates suggest the distribution of health in the population has a potentially important, although not central, role in explaining income inequality. But one could equally claim that variation in economic circumstances is key to the explanation of the health distribution. Indeed, income differences explain almost half of the inequality in predicted health (SAH) in the US.⁷

While descriptive multivariate analysis and decompositions are useful in determining the strength of relationships, they tell us nothing about causality. Nevertheless, we hope to have convinced the reader that the association between income and health is sufficiently strong and pervasive such that it deserves to be probed by those seeking better understanding of the distribution of income, and by others aiming to account for disparities in health. In the next two sections, in turn, we consider the impact that health has on the distribution of income and the effect that income has on the distribution of health.

3. Health determination of economic inequality

3.1 Overview

How might the distribution of health determine the distribution of income? The most obvious effect is through physical and mental capacity for work. We begin this section by considering how health may impact on productivity and wages. We then turn to the relationship between health and the quantity of work. Labour supply may be reduced at both the extensive margin, with illness in middle age tilting the balance in favour of early retirement, and the intensive

---

⁷ Of course, income would explain a much smaller proportion of the greater inequality in actual health.
margin, with part-time work becoming a more attractive proposition for some with a debilitating illness. In high-income countries, the employment effect on earnings will be directly cushioned by disability insurance but resultant moral hazard will indirectly contribute to the effect. Earnings losses may be exacerbated by discrimination but legislation designed to prevent this may increase the impact on employment by constraining wage flexibility. Labour supply may be constrained by current sickness and influenced over the longer term by any downward revision of life expectancy following the onset of a major illness.

In addition to the immediate impact of ill-health on earnings, there may be important lifetime effects operating through education, occupation and marriage. Illness in childhood can constrain opportunities for education and its efficiency in producing knowledge and skills. With few education qualifications, a frail young adult will be limited in his or her choice of occupation. The less healthy may also be constrained in their choice of partner. If there is sorting, such that the least healthy marry others of less than average health, or remain unmarried, then the contribution of health to inequality in household incomes will be even greater than its contribution to inequality in individual incomes. Ill-health may not only directly constrain the work effort of the disabled person but may also indirectly affect the labour supply of the spouse, who may face conflicting demands on his or her time to both replace lost earnings and provide informal care.

Health may impact on the distribution of wealth, both directly and indirectly through income and the accumulation of savings. Since the latter effect is cumulative, the contribution of health to inequality in wealth should be even greater than that to inequality in incomes. The healthy, expecting to live for longer, may save a larger fraction of higher incomes, further increasing wealth disparities by health. When health insurance is incomplete and medical care must be paid for from own resources, illness can directly deplete wealth.
In the subsequent sub-sections, we consider potential effects of health on income through wages, work, human capital, occupation and household formation/spousal earnings. In each case, we elucidate the potential mechanisms and evaluate the evidence relevant to establishing the extent to which health differences contribute to economic inequality.

### 3.2 Health and wages

Individuals in poor health have substantially lower earnings than those in good health. Figure 18.3 shows mean gross earnings of employees in the bottom two categories of SAH as a percentage of the mean in the two top categories for China, the Netherlands and the US estimated from the same data sources as in section 2. While there are discrepancies at the youngest and oldest age groups mainly due to the small sample of individuals working and in less than good health in the Netherlands at these ages, among those working between the ages of 30 and 60, the earnings of those in the worst health are 15-40% below those in the best health in the three countries. In the middle ages, the health gradient in earnings is largest in the Netherlands, which may reflect qualification for partial disability insurance that acts as an earnings subsidy. The relative earnings of the least healthy individuals decline most rapidly with age in China, where lower pension coverage leaves many with little option but to continue working despite deteriorating earnings capacity.

![Figure 18.3: Mean earnings in bottom two categories of self-assessed health as percentage of mean in the top two categories - China, the Netherlands and the US](image)

**Notes:** Authors calculations from CHNS 2006 (China), LISS 2011 (Netherlands), ALP 2008 (US). Samples restricted to those in work. Earnings include gross earnings/salary income from employment and profits from self-employment in the past year in the Netherlands and the US, and refer to gross wages, including bonuses and subsidies, in the last month in China. Self-assessed health is reported from five categories in NL and US and from four in China. See notes to Figure 18.1.
3.2.1 Productivity

Deterioration in health will often decrease labour market productivity and earnings capacity. But the multidimensionality of health and variation in the capacities and skills required for different occupations are reasons to expect a great deal of heterogeneity in the response of wages to health problems. A mobility-impeding disability obviously constrains the productivity of a manual worker much more than it does that of an office worker. Technology, particularly that which enables remote work and reduces the need for commuting, is making the productivity of workers who draw more on their brains than their brawn even less contingent on physical functioning. But the productivity of such workers is dependent on retaining cognitive functioning and is potentially vulnerable to mental health problems. Estimated effects of health on wages derived from measures of different dimensions of health for individuals with different demographic and occupation characteristics should display substantial variability.

The relationship between productivity and physical health is likely to be strongest in developing countries where there is a preponderance of low skilled manual work.8 Identifying the economic returns to health investments has been the motivation for both macro studies revealing a strong positive correlation between economic growth and initial population health (Barro 2003, Barro 2013)9 and micro studies of the relationship between wages and health across individuals in developing countries (Strauss, Thomas 1998). Our focus is not on the impact of health on average income. Rather, we are interested in the extent to which the distribution of health affects the dispersion in incomes.

---

8 Strauss and Thomas (1998) note that the wage elasticity with respect to height (an indicator of health status) was almost eight times larger in Brazil (in the mid 1970s) than it was in the US (in the early 1990s).
9 Barro (2013) maps an extension of the neoclassical growth model to incorporate health, in addition to human, capital and sets out a resulting research agenda, including examination of how health and health policies may impact on the evolution of income inequality.
In low income settings, the relationship between health and productivity can sustain, exacerbate and, in theory, even generate inequality. The crux of the argument is that health constrains productivity and wages provide the means, through nutrition, to sustain health. A negative shock to either health or the wage can generate a downward spiral into a nutrition-based poverty trap (Dasgupta 1993, Dasgupta 1997). The theory of nutrition-based efficiency wages (Leibenstein 1957, Mirlees 1975, Stiglitz 1976, Bliss, Stern 1978, Dasgupta, Ray 1986) generates the prediction of increased inequality as a result of the interdependence of nutritional status and productivity. As nutrition rises above the critical threshold of physiological sustenance, the marginal increases in productivity rise substantially before diminishing returns eventually set in. This non-convexity results in involuntary unemployment since the savings in labour costs from employing the poorly nourished unemployed at lower wages would be more than offset by the resulting loss in productivity. Individuals with initially fewer assets, who can invest less in health, are more likely to be involuntarily unemployed (Dasgupta, Ray 1986). The poor are more likely to be malnourished and sick, and because they are sicker, they are more likely to be unemployed and so fall into destitution. Inequality is exacerbated.

The relevance of this theory has been challenged by the observation that one of its central assumptions – that poor households are constrained to spend almost all their resources on food – is inconsistent with the observed behaviour of the poor (Banerjee, Duflo 2011). Starvation is not perennially present even in very low income countries (Strauss, Thomas 1998). Ill-health may, however, be a source of a poverty trap and contributor to inequality. For the income and asset poor with little formal or informal insurance options, the loss of productivity and earnings precipitated by illness results in reduced consumption opportunities and consequently nutritional deprivation, which further harms health and constrains productivity. The economic impact of illness is likely to be greatest on poor individuals since
their livelihood rests most on their health and because they have fewer assets that can be used to protect consumption and maintain nutritional status when illness strikes. Even if health shocks were evenly distributed across the population, their differential effects would increase economic inequality.

There is a vast body of evidence concerning the impact of health and nutrition on productivity and wages in low and middle income countries. The main motivation for this research is to evaluate the case for investment in health and nutrition programmes as an instrument of development policy. Reviewing this literature would take us well beyond the scope of this chapter. In any case, a number of reviews already exist (Thomas, Frenkenberg 2002, Deolalikar 1988, Strauss, Thomas 1998, Schultz 2005, Commission on Macroeconomics and Health 2001, Schultz 2010). Strauss and Thomas (1998) conclude that there is no robust, consistent evidence from non-experimental studies that ill-health reduces productivity and wages, although it does reduce labour supply. In interpreting this conclusion one needs to bear in mind that the wage response to ill-health can only be studied among employees. This misses the large informal sector of the economy in which the productivity of self-employed, mainly agricultural, workers could be expected to depend on health (Dasgupta 1997). Strauss and Thomas (1998) are convinced of the positive impact of nutritional status (height and body size) on wages and micronutrients (particularly iron) clearly raise productivity. Calorie intake, when accurately measured, is found to have a positive effect on wages, at least among those initially malnourished.

3.2.2 Discrimination
Not all disabilities impede productivity; at least, not in all occupations. Nonetheless, equally productive disabled individuals may be paid less than their able-bodied counterparts because they are perceived to be less productive, or simply because of prejudice.
Discrimination against the disabled, as against other minority groups, comes in two varieties. What economists refers to as *taste discrimination*, and would be more commonly recognized as prejudice, arises from a preference of employers, or other employees, to keep disabled workers at a distance. Since Becker (1957), this has been modeled as a marginal cost, on top of the wage, that a prejudiced employer incurs in employing a member of the minority group. Such an employer will only hire a disabled person at a wage below his or her marginal product. Whether this discriminatory behaviour is sustainable in a competitive market depends on the prevalence of prejudice relative to the supply of disabled labour (Becker 1957). Under competitive conditions, non-prejudiced firms can undercut their prejudiced rivals and discrimination will be competed away (Cain 1986), unless prejudice arises from customers (Kahn 1991).

This model was developed mainly with the purpose of explaining and understanding the consequences of discrimination against ethnic minorities. While some disabilities, or rather handicaps, may still carry a social stigma, most are unlikely to make others, or at least a majority of others, uncomfortable. Stereotyping, or *statistical discrimination* (Phelps 1972, Arrow 1973, Aigner, Cain 1977), seems a more probable source of bias against disabled individuals. In making appointments and wage offers, it is optimal for an employer to supplement information obtained from a noisy signal of productivity, such as a test score or qualifications, with knowledge of the average productivity of a group to which the applicant is observed to belong. Presuming disabled individuals are, on average, less productive, a disabled person would be offered a lower wage than a non-disabled applicant who performed no better with respect to the assessment criteria.

This theory does not help us explain earnings differentials between the disabled and non-disabled over and above those attributable to productivity differences. But it can explain part of the reason for the productivity deficit. If test scores, or qualifications, are a noisier signal of
productivity for the disabled, perhaps because the tests are designed to discriminate between able-bodied applicants, then employers will put less weight on these criteria and more on the observed disability. Faced with a lower return, this group will invest less in human capital. Inequality will be greater than it would be if employers where blind to disability status, or legislation successfully forced them to act as if they were.

Besides its inability to explain earnings differentials beyond those attributable to productivity, the relevance of statistical discrimination as an explanation for health-related wage differences depends upon the extent to which these differences exist across easily recognizable disabilities with known average productivity differentials. A blind man is easily recognized and categorized. Someone with a heart condition is not. Even if all health conditions were observable, perhaps because applicants are required to declare them, how much do employers know about even average productivity specific to them? Rather than responding to an immediately recognizable disability group, it may be that employers have only a partial, perhaps unconsciously biased, understanding of the productivity implication of an incompletely comprehensible health condition.

Empirical identification of discrimination against the disabled is difficult since disability, being an impairment of functioning, will certainly reduce productivity in many jobs. Getting hold of data that make it possible to control for real differences in productivity, and so isolate wage differences attributable to discrimination, is a tall order. Studies that control for little or no differences in health (Kidd, Sloane et al. 2000) cannot credibly claim to identify discrimination against the disabled. But controlling for impaired functioning while comparing wage differences between the disabled and non-disabled seems to be like asking to have one’s cake and eat it. One approach is to concentrate on the wage difference between individuals with easily observed disabilities, such as blindness, paralysis or loss of limb, which may be more likely to evoke prejudice and others who may be completely able-bodied or disabled by
a condition, such as a back pain or heart condition, that is not observable. US data from the 1970s and 1980s reveal that one-third to one-half of the wage differential between these groups is unexplained by wage determinants, including a battery of health indicators intended to capture differences in functional impairments (Johnson, Lambrinos 1985, Baldwin, Johnson 1994). This finding is indicative of substantial discrimination only if the controls are sufficient to mop up any productivity differentials. DeLeire (2001) suggests another approach which involves assuming there are no unobserved productivity differences between individuals who report a health problem but no work limitation arising from this and others reporting no health problem. In that case, all of the wage difference between these two groups can be attributed to discrimination. This can be taken as indicative of the discrimination against those with a work limiting disability under the further assumption that the degree of discrimination is independent of the productivity loss arising from disability. Under these assumptions, only 7-11% of the wage difference not explained by observable characteristics between US males with a work limiting disability and those with no disability could be attributed to discrimination. It is difficult to hazard a guess as to whether this estimate lies closer to the truth than the larger earlier one given that both rest on rather strong assumptions.

3.2.3 Non-wage costs and non-pecuniary benefits

The previous two sub-sections considered wage variation arising from health-related differences in (perceived) productivity. The employer was assumed powerless to correct productivity differentials. A richer model of the demand for disabled labour relaxes this assumption. The productivity of someone bound to a wheelchair is contingent on adjustments made to the workplace - ramps, elevators, adjustable desk, etc. Installation of such facilities involves incurring a fixed cost that pays off through raising the marginal product of disabled workers (Acemoglu, Angrist 2001). Treating labour as a quasi-fixed factor, with the

---

simplifying assumption that there are fixed costs of employing disabled but not able-bodied workers, if after workplace modifications the productivity gap between disabled and able-bodied workers is closed, the wage paid to the former will be lower by the amount of the (discounted) fixed costs (Acemoglu, Angrist 2001). Wage differentials need not reflect only productivity differences or discrimination. Even with perfect measures of productivity, an empirical test of discriminatory behaviour would be difficult. A second implication of this model is that employers are likely to be particularly apprehensive about appointing disabled workers. Fixed costs incurred up-front must be compared with expectations of future marginal products and wages. A risk-averse employer will opt for labour with a higher proportion of variable costs.

Health-related wage variation could also arise from a willingness of employees to trade wage for non-wage benefits. The onset of a chronic condition would be expected to increase the value attached to employer-provided health insurance (Currie, Madrian 1999). A worker suffering from a long-term illness would be more likely to accept a wage cut, or to forgo a pay rise, for fear of not being able to obtain insurance in a better paying job. Note that such health-related wage differentials do not imply differences in well-being. The individual is choosing to accept a lower price for his labour in return for obtaining a lower price at which he can purchase health insurance. Nevertheless, this would be an additional mechanism through which health differences may contribute to inequality in measured income, at last in countries with employment-based health insurance.

3.2.4 Evidence

Theory identifies mechanisms through which ill-health may reduce wages. But how large is the effect? Is health-induced variation large or small relative to overall wage inequality? The usual econometric demons – selection, omitted variables, reverse causality and measurement error – hinder attempts to answer these empirical questions. Evaluation of the evidence
largely comes down to assessing the extent to which these problems have been overcome, or avoided. Ill-health is likely to be a major reason for labour force withdrawal (see next section). Estimation of the impact of health on wages from a cross-section of workers or a balanced panel of individuals in continuous employment will overlook those whose wage opportunities were reduced most by ill-health and decided to stop working. Correction of this selection bias requires modeling employment, in addition to wages, with health allowed to impact on both. With panel data, there may also be health related attrition: those experiencing a marked deterioration in health being more likely to drop out of the sample. Recognizing that individuals can influence their health through lifestyle, for example, leads to realization that the same unobservable factors, such as time preferences, risk attitudes and schooling quality, that influence job choices and so wages may also condition investments in health. If panel data are available, then differencing can be used to purge the time invariant unobservables correlated with health, or efficiency gains may be sought by using averages of assumed exogenous time varying covariates to instrument health (Hausman, Taylor 1981). Neither of these solutions is sufficient to remove bias if there is direct dependence of health on the wage. This is implied by Grossman’s seminal model of health determination (Grossman 1972a, Grossman 1972b), according to which the wage influences both the costs of, and the returns to, investments in health (see section 4.1). Correcting or avoiding the threat of simultaneity bias requires identification from exogenous variation in health that does not arise from wage differences. Prices of medical care and, in a developing country context, the local disease environment have been used as instruments for health (Strauss, Thomas 1998). However, it can be difficult to find variation in prices that is not endogenous to the choice of medical care provider and geographic variation may be a rather weak instrument (Currie, Madrian 1999). Disease exposure is often correlated with weather and agricultural conditions that would be expected to impact on wages directly.
Currie and Madrian (1999) provide a comprehensive review of the US evidence on the health impact on wages (and on labour supply) up to the turn of the century. They note three main deficiencies in this evidence base. First, estimates are sensitive to the measure of health and variability in the measures adopted impedes comparability across studies. We would add that this sensitivity does not merely reflect inconsistency in the measurement of health but is due to intrinsic heterogeneity in the effect depending on the nature of the health condition. Second, few studies attempted to correct for the potential endogeneity of health and those that did relied on rather dubious exclusion restrictions. Third, most of the evidence available referred to white (US) males. Picking up from where Currie and Madrian (1999) left off, in the remainder of this sub-section we focus on the evidence published since 1999 using data from high-income countries. The latter two criticisms have, to an extent, been addressed in the more recent literature. The increasing availability of panel data, particularly on older populations that experience most variability in health, as well as population level administrative data even more recently, has reduced reliance on instruments to deal with endogeneity. While many studies still tend to focus on males, there are many exceptions and the evidence comes from a wider spread of countries. All studies cited are summarized in Table 18.3 for evidence relating to the US and 18.4 for studies that use data from European and other high income countries.

A fixed effects estimate obtained from retrospective life history data collected in the first wave (1992-3) of the US Health and Retirement Study (HRS) suggests that a work limitation lasting at least three months reduces the wage rate by 4.2% for males and twice that for females aged 50-60 (Pelkowski, Berger 2004). Given that 7-9% of individuals in this age range report such a health condition, these estimates suggest that ill-health makes a substantial, though not dramatic, contribution to wage inequality. Using twenty-five years of longitudinal data from the US Panel Study of Income Dynamics (PSID), Charles (2003)
obtains a fixed effects estimate (corrected for selection into employment) only half as large for men for a similar measure of ill-health experienced by almost one-third of the sample at some time during the panel. This would suggest a much more modest contribution of ill-health to wage inequality. The lower estimate obtained by Charles may be attributable not only to the use of panel, rather than retrospective, data, but also to estimation using a younger sample. An analysis of the same dataset and health measure, and taking account of simultaneity as well as selection and unobservable heterogeneity, finds that below the age of 35 and above the age of 62 there is little difference between the wage profiles of individuals in good and bad health (French 2005). But in the prime ages of working life, individuals with a work limitation can command wages around 8-17% below the wages of those in good health.

Using UK panel data and the Hausman and Taylor (1981) estimator, Contoyannis and Rice (2001) find a significant effect of psychological health, but not general SAH, on wages for males but no interpretation of the magnitude of the effect is given. For females in full time employment, there is a significant effect of SAH; moving from less than good health to excellent health is estimated to result in a rather modest wage increase of less than 3%.

A cross-section analysis of data on Australian men that attempts to allow for full simultaneity finds a large effect of SAH on wages (Cai 2009). But the instruments used (health conditions and behaviour) are of dubious validity and exploitation of the panel dimension of the data using a fixed effects estimator results in a large reduction of the estimate and loss of its significance.11 Jackle and Himmler (2010) also resort to instruments in order to deal with endogeneity arising from more than correlated time invariant unobservables in analysis of German panel data. They assume that past doctor visits determine health but not labour force

11 The author attributes this to greater measurement error in health changes and failure of the fixed effects estimator to deal with correlated idiosyncratic errors. Alternatively, it could be that the IV estimate is upwardly biased by invalid instruments.
participation or wages conditional on this. The rationale is that past medical care is the investment response to previous health shocks and need not be correlated with current labour market outcomes given current health. This may be so but it does not allow for the possibility that individuals visit a doctor to obtain a sick note to justify work absence. For males, a worsening in reported health from *excellent to poor* is estimated to result in a 4.8% drop in the hourly wage. For females, there is no significant effect.

While differences in health indicators and estimators still make it difficult to compare estimates, we tentatively conclude that ill-health does reduce wages in high income economies but the effect is more likely modest than substantial. Since most studies estimate the wage response to ill-health while controlling for occupation, this conclusion refers to the degree to which the wage adjusts within a given job. A larger wage effect may arise through ill-health induced changes in occupation. We examine this effect in section 3.5.

[Tables 18.3 and 18.4 here]

### 3.3 Health and work

As would be expected, employment rates vary a great deal with health. This is illustrated in Figure 18.4 for China, the Netherlands and the US. In the two high income countries, already in young adulthood individuals reporting *poor* health are much less likely to be working than their contemporaries reporting better health. The difference in employment by health widens until middle-age, after which early retirement begins to reduce the labour force participation of even those in good health and the employment gap narrows. The relationship between employment and health is different in China in two respects. First, there is little or no difference in employment by health in young adulthood. Second, while a gap opens up at older ages, it never becomes as wide as that observed in the Netherlands and the US. This is partly an artefact of self-assessed health being reported in four, rather than five, categories in
the Chinese survey. But this is not the whole story. In the US around the age of 50, the employment rate difference between those reporting fair health and those reporting excellent or very good health is about twice as large as the difference between those reporting poor health and those reporting very good health in China at the same age. The relationship between employment and health is weaker in China. This is not what one would expect given the differences in the structures of the economies and the greater role of manual labour in China. It may be that the more generous social protection in the high income countries allows individuals experiencing health problems to more easily withdraw from the labour market.

In this section we consider a number of mechanisms through which health may impact on employment, including the incentive effects of disability insurance.

![Figure 18.4: Employment rates by self-assessed health and age in China, the Netherlands and the US](image)

**Notes:** Authors calculations from 2006 CHNS (China), 2011 LISS (NL) and 2008 ALP (US). Employment includes full and part time work. For details of SAH by country see notes to Figure 18.1.

### 3.3.1 **Incapacity and involuntary unemployment**

The impact of ill-health on work may seem obvious. If you are sick, you cannot work. For relatively short term, acute illnesses this is a reasonably adequate description of the effect. But it is an effect that could only explain temporary interruptions to earnings and income. Any substantial contribution of the distribution of health to the distribution of income is unlikely to operate through short term sickness. A few chronic medical conditions are completely
incapacitating. But most reduce capacity for work to some degree. Ill-health may reduce productivity but it is unlikely that the marginal product is pushed to zero in all possible jobs. The wage could fall below the level at which work is considered worthwhile, but that is a choice rather than a fait accompli.

This reasoning rests on the assumption that wages are perfectly flexible. Evidence of a moderate impact of health on wages (see section 3.2.5) may reflect institutional constraints on wage flexibility. Equal pay and anti-discrimination laws typically make it illegal for employers to pay disabled workers less than their able-bodied colleagues doing the same job. This may succeed in constraining health-related wage inequality but exacerbate disparities in employment.

The 1990 Americans with Disabilities Act (ADA) compels employers to accommodate disabled workers through adjustments to the workplace and outlaws discrimination against the disabled in hiring, firing and pay. In theory, the impact on employment of the disabled is ambiguous. The threat of legal action for discrimination in hiring would tend to increase employment, while increased accommodation and hiring costs would reduce employment. Acemoglu and Angrist (2001) argue that the negative effects are likely to dominate. Before and after legislation trends in employment are consistent with this prediction (DeLeire 2000, Acemoglu, Angrist 2001). Employment of disabled individuals was also reduced immediately after the introduction of the UK Disability Discrimination Act in 1996, which imposed similar obligations on employers as the ADA, before it recovered somewhat (Bell, Heitmueller 2009). Digging deeper into the effect in the US, Hotchkiss (2004) reveals that it is not due to individuals always classified as disabled being more likely to leave employment or less likely to enter employment. Rather, it is due to nonparticipants in the labour market reclassifying themselves as disabled after the passing of the legislation. Whether antidiscrimination legislation makes it more difficult for disabled individuals to obtain work may still be an open
question. What seems clear is that there is no evidence that major legislation makes it easier for disabled individuals to gain employment.

3.3.2 Disability insurance

Any illness induced reduction in the offer wage makes labour force participation less financially attractive. If the decline in health is sufficient to qualify for disability insurance, then financial disincentives to work are compounded. Qualification for disability insurance is not unambiguous (Diamond, Sheshinski 1995). It is typically not determined by the presence of a precisely defined medical condition but is assessed on the basis of the vague concept of ‘capability of performing paid work’, perhaps taking account of workplace conditions and occupation. There is subjectivity in whether a person considers himself incapable of work, as well as whether the adjudication officer agrees. Financial incentives can tilt the balance in favour of applying for disability insurance. For a given degree of work incapacity, withdrawal from employment is more likely when social protection is available to cushion the resulting loss of income (Parsons 1980, Bound, Burkhauser 1999, Gruber 2000, Autor, Duggan 2006).

By increasing the likelihood of labour force withdrawal but compensating for the resulting income loss, the existence and generosity of disability insurance (DI) may simultaneously strengthen the relationship between health and earnings, and weaken the relationship between health and income.12 For a given distribution of health, income would be expected to be more equally distributed in countries with generous DI. But the equalizing effect of social protection may be weakened by a moral hazard effect that is stronger for individuals with lower earnings potential. If, as is usually the case, the replacement rate is decreasing with pre-disability earnings, the financial disincentive to continue in employment will be greatest for

12 Hurd and Kapteyn (2003) find that income change is more sensitive to the level of self-assessed health of individuals aged 51-64 in the Netherlands, which has more generous disability insurance (DI), than in the US. A possible explanation is that the moral hazard effect outweighs the income replacement effect of DI. Individuals in poor health are more likely to withdraw from employment in the Netherlands, consistent with what is observed in Figure 18.4 at prime working ages.
lower paid workers. As a result of these differential incentives, onset of a given disability is more likely to result in employment withdrawal and loss of earnings (only partially replaced by DI) of lower paid workers. This differential moral hazard effect will tend to increase income inequality relative to the hypothetical situation in which the financial incentives arising from DI have no impact on employment. But in the complete absence of DI and the income protection it provides, income inequality would be likely to be even greater. In addition to the strength of the differential moral hazard effect, the extent to which DI reduces income inequality will depend on the incidence of disability. The equalizing effect will be greater if, as is likely, the poor are more likely to become disabled. The lower paid are more likely to both benefit from social disability insurance and respond to its financial disincentives to work.

Disability insurance rolls have been rising over much of the past 30 years in many high-income countries (OECD 2010, Wise 2012). Steep downward trends in mortality rates suggest that this is not because populations are becoming less healthy. Looser eligibility criteria and increased returns to claiming DI relative to those available from work are considered to be the chief culprits (Bound, Burkhauser 1999, Autor, Duggan 2006, OECD 2010, Wise 2012). We will not evaluate the evidence concerning the strength of the moral hazard effect of disability insurance (Bound, Burkhauser 1999) but concentrate on the proposition that it interacts with increased economic inequality arising from structural changes in the economy to reduce labour force participation of low skilled, low paid workers (Autor, Duggan 2006, Autor, Duggan 2003).

The falling relative wages and employment opportunities experienced by low skilled workers in the US and other high-income countries in recent decades increases the attractiveness of DI for this group (Black, Kermit et al. 2002, Autor, Duggan 2003). Dependence of DI awards on ability to engage in gainful employment results in DI applications tending to rise in an
economic downturn (Autor, Duggan 2006) and suspicion that some governments deliberately use DI to disguise long term unemployment. After loosening of the Social Security Disability Insurance (SSDI) eligibility criteria in the US in the mid-1980s, the sensitivity of applications to adverse economic conditions increased by at least twofold and high school dropouts became twice as likely to exit the labour force on occurrence of a negative shock to the economy (Autor, Duggan 2003). The effect is compounded by the indexing of the SSDI benefit formula to average wage growth. As a result, the replacement rate has increased for individuals whose wage growth lagged the average, as has been the case for low paid workers in the US over the past 30 years (Autor, Duggan 2006, Autor, Duggan 2003). The replacement rate was further increased by rising real expenditures on Medicare, the health insurance programme for the elderly that SSDI beneficiaries are given entitlement to before reaching the age of 65. Taking into account these fringe benefits, the DI replacement rate for a 50-61 year old male at the 10\textsuperscript{th} percentile of the earnings distribution increased from 68\% in 1984 to 86\% in 2002 (Autor, Duggan 2006). At the 90\textsuperscript{th} percentile, the increase was much more modest, from 18\% to 22\%. As would be expected given these differential incentives, SSDI enrolment rates are much higher, and have increased much more rapidly, for low skilled individuals. For male high school dropouts aged 55-64, the rate increased by five percentage points between 1984 and 2004 to reach almost 20\% (Autor, Duggan 2006). The increase was only one percentage point (to reach 3.7\%) for the college educated.

In the US, reduced earnings prospects relative to the average and increased DI replacement income relative to those reduced earnings, separately and in combination, have reduced the incentives for low skilled workers with a health problem to continue working. Increased reliance of this disadvantaged group on DI may be both a consequence of rising economic inequality and, given the replacement ratio is less than one, a contributor to it. This process
may also operate in Europe, which also witnessed increased wage inequality and, at times, rising DI rolls over recent decades.

3.3.3 Preferences
In addition to an indirect effect through disability insurance, ill-health may shift the reservation wage directly by changing preferences for consumption relative to leisure. The direction of the effect is ambiguous. Ill-health would be expected to increase the disutility of work. But it may also reduce the marginal utility of a number of leisure activities, e.g. sport. The direction of the effect on the marginal utility of consumption is even more difficult to predict. A disabled person may derive less, or no, pleasure from some goods, e.g. sports equipment, travel, but become more dependent on others, e.g. pharmaceuticals, heating, private as opposed to public transport. Comparing the relationship between subjective well-being and consumption – proxied by permanent income – for older Americans with and without chronic illness, Finkelstein, Luttmer et al (forthcoming) infer that the marginal utility of consumption decreases with ill-health. If we assume that the increased disutility of work dominates, such that the marginal utility of leisure rises, this gives a clear prediction that the marginal rate of substitution of leisure for consumption rises when health falls. The reservation wage rises and labour force withdrawal becomes more likely. More direct evidence suggests that a work-limiting health problem is equivalent to ageing around four years in the extent to which it increases the willingness of older (58+) Americans to trade consumption for leisure (Gustman, Steinmeier 1986, Gustman, Steinmeier 1986).

While these estimates support the plausible proposition that ill-health shifts preferences away from work, one should be careful not to overlook the multidimensionality of health. A

---

13 The evidence from Finkelstein, Luttmer et al (forthcoming) is obtained from a sample of non-working, elderly individuals and so one cannot necessarily infer from this how ill-health affects preferences for consumption relative to leisure.

14 For other estimates of health-specific utility function parameters in models of health and retirement, see (Sickles, Yazbeck 1998, Bound, Stinebrickner et al. 2010).
physical disability may reduce the marginal utility of many leisure pastimes as much as, or
more than, it raises the disutility of work. A chronic illness such as diabetes may have little or
no impact on preferences for leisure relative to consumption. The empirical content of a
prediction that preferences for work increase with a characteristic we conveniently refer to as
health is blurred if that characteristic, or at least the means of measuring it, is not well-
defined. We are rather attached to the concept of health in the social sciences despite finding
it difficult to define what we mean by it (Twaddle 1974). Estimates of the impact of ill-health
on work-leisure preferences are likely to vary with the dimensions of health examined and the
indicators used to measure them.

3.3.4 Life expectancy

Expected longevity is an additional mechanism through which differences in health may
contribute to observed differences in income and wealth in a cross-section. In the standard
life-cycle model of consumption with no bequest motive in which there is dissaving before
death, a longer length of life is predicted to increase labour supply (and saving) at any given
age (Hamermesh 1984, Chang 1991). This is basically a wealth effect. Increased lifespan
implies greater lifetime potential income. The resulting increase in demand for consumption
prompts a rise in labour supply.15 We may refer to this as the horizon effect. Individuals in
poor health work less because they do not have to provide for an extended old age. For a
given degree of functional impairment, illnesses that are life threatening, or at least
shortening, should be observed to reduce earnings by more than chronic disabilities that
present no threat to longevity (McClellan 1998).

Recognizing that the length of life is uncertain and annuity markets are incomplete gives rise
to an offsetting effect. Through the mortality risk on savings i.e. the prospect of dying before

---

15 Similarly, increased loss of time due to sickness within any period of life is predicted to reduce labour supply
in that period, although increases in the expectation and uncertainty of future sickness are predicted to increase
current work effort (O'Donnell 1995).
being able to enjoy the fruits of one’s savings, variation in the survival probability affects the marginal return on savings and, consequently, the marginal rate of substitution between consumption and leisure (Chang 1991, Kalemli-Ozcan, Weil 2010). Through this uncertainty effect, a reduction in the probability of death raises the return on and so level of savings, making it possible to reduce labour supply, perhaps by retiring earlier, even when there is a longer expected length of life to be provided for. This uncertainty effect is less likely to dominate the horizon effect the lower is the initial mortality rate (Kalemli-Ozcan, Weil 2010).

In high income countries, one expects ill-health and reduced longevity to be associated with lower earnings (and wealth).

### 3.3.5 Evidence

Apart from the ambiguous effects through preferences and longevity, theory gives the clear prediction that ill-health reduces work effort. Under reasonable assumptions for high-income economies, the effects through preferences and longevity go in the same direction. But what is the size of the overall effect of ill-health on employment?16

Most studies address this question using samples of older individuals. Effectively, they estimate the impact of health on (early) retirement. In general, health is found to be an important determinant of retirement (Currie, Madrian 1999, Lindeboom 2012). How important depends on the measure of health and estimator adopted, as well as the context.

Estimating the effect of health on employment is complicated by similar econometric obstacles to those confronted when trying to identify the impact of health on wages, only the measurement error problem becomes particularly thorny. Most studies have relied on survey data and self-reported measures of ill-health. One would expect the reporting of health to be

---

16 We address this question in the context of high income countries. In low income countries with less formal labour markets, identification of the impact of health on employment and earnings is more challenging. The literature tends to focus on the extent to which households can smooth consumption over health shocks in the absence of formal health and disability insurance (Townsend 1994, Gertler, Gruber 2001, Mohanan forthcoming). Reviewing that literature would take us beyond the scope of our objectives.
endogenous to employment. Put most crudely, individuals who have decided not to work may lie about their capacity to work either to reduce stigma attached to voluntary inactivity or because they are claiming disability insurance. But the phenomenon need not be so blatant. The threshold of functioning at which an individual considers himself to be incapable of work may be influenced, possibly subconsciously, by financial incentives to work, job stimulus, length of working life, contact with others claiming disability insurance, etc. Reported ill-health may reflect motivations for not working. This so-called *justification bias* has been a major concern in the literature (Bound 1991). The evidence tends to suggest that it leads to substantial overestimation of the effect of ill-health on employment (Bazzoli 1985, Lindeboom, Kerkhofs 2009, Bound, Stinebrickner et al. 2010, Lindeboom 2012), although there are dissenting findings (Stern 1989, Dwyer, Mitchell 1999).

Replacing reported work incapacity with more objective indicators of chronic illnesses or future mortality reduces the risk of justification bias but increases classical measurement error and may result in underestimation of the effect (Bound 1991). Instrumenting reported work limitations with more objective health indicators is arguably a better approach (Stern 1989, Bound 1991, Bound, Stinebrickner et al. 2010) but requires that the indicators are free of the justification bias, which may be a strong assumption given these are often also self-reported.

All studies from North America and Europe that attempt to deal with justification bias by treating self-reported health as endogenous confirm that health is an important determinant of labour force participation (Sickles, Taubman 1986, Stern 1989, Bound, Schoenbaum et al. 1999, Dwyer, Mitchell 1999, Kerkhofs, Lindeboom et al. 1999, Blau, Gilleskie 2001, Au, Crossley et al. 2005, Disney, Emmerson et al. 2006, Lindeboom, Kerkhofs 2009, Jones, Rice et al. 2010, Brown, Roberts et al. 2010).¹⁷ Health also emerges as a strong determinant of

---

¹⁷ All studies cited in this section that are not included in the tables presented in Currie and Madrian (1999) are summarised in Table 18.3 for US and Table 18.4 for elsewhere.
retirement in structural life cycle models of older US males (Sickles, Yazbeck 1998, French 2005, Bound, Stinebrickner et al. 2010). Bound, Stinebrickner et al (2010) find that while dealing with justification bias by instrumenting self-reported health with objective indicators greatly reduces the health effect, it remains very large. Before the early retirement age, an older single male in bad health is five times more likely to withdraw from the labour force than an equivalent in good health. On reaching 62, the age at which the Social Security pension can first be claimed in the US and when financial incentives shift in favour of retirement, the probability of withdrawal rises from 0.1 to 0.17 for those in poor health and from 0.025 to 0.049 for those in good health. In absolute terms, those in poor health react more to the financial incentives such that the difference in employment probabilities widens when it becomes financially more advantageous to retire. Ill-health and financial incentives interact. Financial incentives appear to tilt the balance toward considering a health problem to be incapacitating. From a low level of health, marginal deteriorations in health have a large effect. At age 62, one-half of a standard deviation decrease in health for someone in poor health raises the probability of labour force exit from 0.17 to 0.27; ten times larger than the change in probability for someone in average health.

Extending the scope of analysis beyond older US males, French (2005) finds a very strong health effect that varies across the life cycle. Below the age of 40, there is no difference in the employment rate of men in good and bad health. At 40, an incapacitating physical or nervous condition is estimated to reduce the employment probability by five percentage points for a stereotypical male. The effect reaches a peak of 60 points at age 58 before declining to 45 points at age 62 and 20 points at age 66. In responding to these very large effects, one needs to bear in mind that they are not corrected for justification bias. Despite the very large estimated health effect, the author argues that it is modest in terms of the proportion of the

---

18 Health is modelled as a latent variable. ‘Good health’ refers to an average score and ‘bad health’ to a score one standard deviation below the average.
total decline in older male labour force participation that can be attributed to population health. The percentage of males reporting an incapacity rises from 20% at age 55 to 37% at age 70. Applying the estimated health effect, this decline in health can explain only 7 of the 74 percentage point fall in labour force participation between these ages. This may be so, but the estimates of both French (2005) and Bound et al (2010) imply that ill-health can account for a substantial fraction of labour force withdrawal before the early retirement age of 62.

Establishing a substantial effect of health on employment only takes us part way toward our objective of gauging the contribution of health to the distribution of income. We need to know the earnings and, ultimately, income consequences of illness-induced loss of employment. Reviewing the earlier US evidence, Currie and Madrian (1999) conclude that ill-health has a large negative effect on earnings, which operates mainly through reduced hours of work, including non-participation, rather than reduced wages.

From the US HRS life history data, Pelkowski and Berger (2004) find that 7-9% of individuals over the age of 50 have experienced a work limitation at some time that, on average, is estimated to result in the loss of half of potential future earnings. By far the greatest part of this substantial loss is from the reduced likelihood of working, which falls to around half of its counterfactual value. Taken at face value, these estimates suggest that health is a major determinant of earnings differences. Caution should be exercised in their interpretation, however. A major concern is the retrospective nature of the data, which may be vulnerable to recall, as well as justification, bias. There could be a tendency to report health events that did have labour market consequences, and to recall changes in labour outcomes that coincided with periods of illness.

Charles (2003), using the US PSID, finds that annual earnings of initially employed individuals fall, on average, by 15% around the time of onset of disability. Given that almost
one-third of the panel experiences a work limitation at least once, this suggests that ill-health is an important contributor to cross-sectional inequality in earnings. But earnings do not remain as depressed as they are immediately following the onset of ill-health. Some illnesses recede and disabilities can be adapted to through re-training. Within two years of the onset of disability about half of the earnings loss is recovered. Subsequently, earnings continue to trend upward, rather slowly, towards the level at which they would have been without having experienced the disability. These findings suggest that Pelkowski and Berger (2004) substantially overestimate lifetime earnings losses by extrapolating from the contemporaneous impact of ill-health on earnings.

While the PSID study provides valuable insight into the impact of ill-health on earnings, it is weakened by reliance on self-reported work limitation with no correction for justification bias. The steepest drops in earnings occur in the period between two and one year prior to the reporting of a work limitation. One interpretation is that health is declining and impeding labour outcomes prior to the point at which a health problem is reported. But another is that decreasing returns to and motivation for work lead to the reporting of a disability, perhaps to justify entry to disability insurance.

Besides its vulnerability to justification bias, another limitation of the reported work capacity measure is that it focuses on functional impairment and does not discriminate other dimensions of health, such as longevity. This is not sufficiently recognized in part of the literature, which tends to presume that the ideal measure of health would be one that accurately informs of work capacity (Bazzoli 1985, Bound 1991, Lindeboom 2012). Implicitly, the assumption is that there is only one mechanism through which health impacts on employment – physical and mental capacity to perform work related tasks. Recognizing the other mechanisms identified in the preceding sub-sections, health conditions that differ
with respect to their implications for current functional impairment, prospects of recovery, survival chances etc can be expected to have differential impacts on work.

McClellan (1998) makes an interesting distinction between three types of health outcomes. *Major health events*, such as a severe heart attack or a stroke, imply both acute and long-term functional impairment and reduced life expectancy. Onset of a *chronic illness*, such as a heart condition or diabetes, does not dramatically affect current functioning but may affect labour supply through expectations given prospects of a degenerative disease and long-term impairment. At the other extreme, *accidents* have an immediate impact on functioning but are less likely to have any effect on preferences or health expectations in the long term. Using the first two waves of the US HRS, McClellan finds evidence consistent with the nature of the health event having an independent effect on employment over and above that of the degree of functional limitation arising from it. For a given change in functioning (measured by Activities of Daily Living (ADL)), the reduction in the employment probability is 40 points greater when it is precipitated by a major health event. The employment probability falls by 14 and 35 percentage points for males incurring a new chronic illness with moderate and major reductions in functioning respectively. Accidents are not significantly associated with a decline in employment. McClellan interprets this as indicative of employment effects being muted when the health event does not have consequences for long term health expectations.\(^\text{19}\)

Smith (2004) extends this type of analysis by using more waves of the HRS to look at longer term effects. Onset of a major health condition (i.e. cancer, heart disease, lung disease) is associated with an immediate reduction of 15 percentage points in the employment probability of males among whom just over half were working at baseline. After eight years, those contracting such an illness are 27 points less likely to be working. The short and long

\(^{19}\) It could also be that the number of survey observations experiencing an accident does not provide sufficient power to precisely estimate the effect. Using population data, Möller Dano (2005) and Halla et al (2011) do find significant effects of accidents on employment (see below).
term effects for those succumbing to a minor chronic illness (i.e. hypertension, diabetes, arthritis, heart attack, angina and stroke) are reductions of 4 and 11 percentage points respectively. The average loss in annual household income immediately following the onset of a major chronic illness is $4000, which rises to $6250 after eight years. Cumulating these losses gives a total loss of income over eight years of $37,000. The cumulative income loss associated with a minor condition is almost $9000.

This evidence of sustained reductions in employment and income appears to contradict that of Charles (2003). It does not once the older age of the HRS sample is recognized and the heterogeneity in the effects estimated by Charles is scrutinized. Men that become chronically disabled, defined as reporting a work limitation in every period after the initial onset, experience an estimated initial earnings loss of 21% with little or no recovery over time. The initial loss is also greater and the recovery absent for older men. So, both studies find substantial permanent losses of income for older males succumbing to a chronic condition.

Bound, Schoenbaum et al (1999) focus on the implications of the dynamic evolution of health for the continued labour force participation of older (50+) US workers. They find that it is health deterioration, rather than the health level arrived at, that most affects labour force exit. Employment is not simply dependent on current functional capacity since there can be adaptation to impairments. A drop to any given level of health implies lower expectations of future health that may have an independent effect on the decision to continue in employment.

Disney, Emmerson et al (2006) take the same approach with British data and find a different pattern of behaviour. Controlling for past health, lower current health still increases the likelihood of job exit. But for any given level of current health, a lower level of past health also raises the probability of retirement. A possible interpretation of this result is that individuals take time to revise their health expectations downward. Only when health is persistently lower is the decision made to retire, which may be difficult to revoke.
Cross-country comparisons can be useful for exploring the extent to which the employment and income responses to ill-health appear to be influenced by employment and social policies. García-Gómez (2011) compares the likelihood of continuing in employment following a sustained drop in SAH in nine EU countries. She finds the employment effects are largest in Ireland, a country in which disability insurance claimants are not allowed to engage in any form of paid work, followed by Denmark and the Netherlands, where replacement rates are highest and, as in Ireland, there is no quota on the percentage of employees that must be registered disabled. In France and Italy, the two countries that impose the highest quotas on disabled employees, there is no significant impact on employment. These findings suggest, but do not confirm, that the employment (and income) effect of ill-health is highly contingent on policies influencing both the demand and supply of disabled labour.

Rather than attempt to identify the health effect from data on self-reported general health, which is difficult to interpret and potentially endogenous, or reported work incapacity, which is possibly even more endogenous due to justification bias, three European studies have concentrated on more narrowly defined health events that are abrupt and unforeseen and so more plausibly exogenous to life cycle planning of health and labour supply (Moller Dano 2005, Halla, Zweimuller 2013, García Gómez, van Kippersluis et al. 2013). Unlike structural models that aim to estimate the endogenously determined life cycle profiles of health and labour supply (Sickles, Yazbeck 1998, French 2005, Bound, Stinebrickner et al. 2010), the objective of these studies is to exploit some unanticipated shift in the health profile to estimate the response of employment and income. Two of the studies use accidents (Halla, Zweimuller 2013, Moller Dano 2005), while the third relies on urgent and unscheduled hospital admissions (García Gómez, van Kippersluis et al. 2013). Implementation of this strategy is made feasible by the availability of population, or near population, data from administrative registers providing sufficient observations of relatively rare health events for which matches
can be found from millions of control observations. Use of administrative records greatly reduces measurement error and avoids the justification bias that plagues estimates based on reported health. All three of the studies combine matching, to deal with observable differences, with taking difference-in-differences to eliminate correlation of the health event with time invariant unobservables.

These studies consistently find that a health shock reduces the probability of employment. The estimates range from a 3.3 percentage point reduction as a result of a commuting accident in Austria (Halla, Zweimuller 2013), through a 7.1 point drop due to an acute hospital admission in the Netherlands (García Gómez et al, 2013), to a 11.8 point fall following a road accident experienced by men in Denmark (Moller Dano 2005). That the estimates differ in magnitude is to be expected given the narrow definition of the health events from which they are identified. Commuting accidents mostly give rise to musculoskeletal impairments, while diseases of the circulatory, digestive and respiratory are all important causes of acute hospital admissions. Such different conditions would be expected to have different effects. Focus on specific health events that occur suddenly enhances the internal validity of these studies, but generalization to other forms of health deterioration cannot be presumed. There is no avoiding the multidimensionality of health and the consequent heterogeneity in its effects.

These studies confirm that ill-health causes employment to fall; a conclusion that could be made with less certitude from more weakly identified estimates. A less predictable finding is that the effect is persistent. All three studies find that the probability of employment remains reduced by a health shock for at least five years following its occurrence. This contrasts with what Charles (2003) finds for prime working age US men. The difference is most likely attributable to the lack of incentives for DI recipients in continental Europe to move off the roll.

---

20 There is no significant impact of a road accident on the employment probability of women.
In Austria and the Netherlands, but not in Denmark, the impact on employment is greater for women. In the same two countries, the effects are greater on older persons and blue collar (Austria) or low income (the Netherlands) workers. This is consistent with evidence from the UK showing that older poor individuals are more likely exit the labour force by entering Disability Insurance (Banks 2006). It is also consistent with US evidence from the PSID that earnings losses from ill-health are larger and more sustained for non-whites and the poorly educated (Charles 2003), characteristics associated with working in industries and occupations in which productivity is more contingent on physical health and that identify low skilled workers for whom, as observed in section 3.3.2, labour market opportunities have deteriorated and dependence on disability insurance has increased. The less privileged may not only be more likely to be struck by ill-health but their employment and incomes are also more contingent on their health. Ill-health may increase economic inequality through both its skewed incidence and its differential effect.

In the Netherlands, an acute hospital admission results in an average reduction of around 5% in personal income two years after the health shock, with little or no recovery over the following four years (García Gómez, van Kippersluis et al. 2013). For individuals who remain in employment, income falls by only 3%, indicative of very modest reductions in wages and hours of work at the intensive margin. Those moving onto disability insurance experience an income loss of one-third, which is broadly consistent with the DI replacement rate. While this is a substantial drop, the moral hazard effects of providing more complete income protection are likely to be large in a country where at one time 10% of the working age population was on DI. But Denmark does offer even greater insurance, with an average 12% average drop in male earnings maintained for six years after a road accident offset by a rise in transfer income such that there is no significant change in total income (Moller Dano 2005). This does not imply that there is complete insurance. Presumably the 12% who lose employment do
experience income losses but this is not reported. For women, there is no significant drop in earnings or loss of income. These findings are consistent with the income consequences of ill-health being muted in a country, like Denmark, with a generous welfare state. But one should keep in mind that it is the effect of a road accident that is estimated. According to McClellan (1998) accidents should have the mildest economic consequences because the induced health change may be temporary and longer term functioning and survival expectations may be little affected. This is the limitation of this approach. While one can be confident that the estimate does accurately capture the effect of the health change studied, the rarity of the narrowly defined event reduces the relevance of the evidence to the broader question of the extent to which the overall variation in health contributes to observed economic inequality in the population.

3.4 Early life health determinants of later life economic inequality

The focus on the income effects of ill-health in adulthood until this point risks missing much of the action. A rapidly growing literature, to which Janet Currie and James Heckman are leading contributors, argues that early life – even pre-birth – and childhood conditions, including health, explain much of the variation in economic outcomes across adults (Cunha, Heckman et al. 2006, Heckman, Stixrud et al. 2006, Heckman 2007, Currie 2009, Almond, Currie 2012). It has been estimated that a staggering 50% of inequality in the present value of lifetime earnings in the US can be explained by factors known at age 18 (Cunha, Heckman 2009). Although most, such as parental occupation, are not directly health related, exposure to health risks in the womb, infancy and childhood are potentially important components of these economically significant conditions. Ill-health in early life may directly constrain health capital in later life and impede the accumulation of non-health human capital. Both effects would reduce earnings potential.
One may distinguish three broad mechanisms through which early life health may impact on economic outcomes in adulthood, differentiated by the life stage in which they become manifest. Nutritional deprivation and exposure to health risks in utero and in infancy can directly impair cognitive functioning and lead to childhood health problems that interfere with the acquisition of cognitive, and possibly non-cognitive, skills. A second route is through education. Ill-health in childhood and adolescence may restrict opportunities to acquire education, and impaired cognitive functioning arising from insults to health in infancy may reduce the efficiency of schooling in producing educational qualifications. The third mechanism operates through health capital, as opposed to other forms of human capital. There may be persistence in ill-health from childhood to adulthood. More dramatically, exposure to health risks in the womb may do lasting physiological damage, which becomes manifest with the onset of disease in middle age. Earnings may subsequently fall, as is clear from the evidence reviewed in the previous section.

In the following subsections we discuss each of these three broad mechanisms by which infant and childhood health may constrain economic success in adulthood. This takes us into territory that has traditionally been the domain of psychology and epidemiology but in which economists are increasingly daring to venture. We end the section by summarizing and evaluating the evidence on the extent to which adult economic outcomes are determined by early life health conditions. We do not provide detailed reviews of the rapidly growing literatures, which have already been provided by researchers with far greater expertise (Currie 2009, Almond, Currie 2011, Almond, Currie 2012, Cunha, Heckman et al. 2006). Our focus is on what the literature has to say about the contribution of health to income inequality and to understanding the association between income and health in adulthood.
3.4.1 Health, cognitive and noncognitive capabilities

There is abundant evidence that cognitive functioning is a strong predictor of wages (Jencks 1979, Herrnstein, Murray 1994, Cawley, Heckman et al. 2001). We are interested in whether cognitive functioning measured during childhood determines economic success later in life and whether childhood cognitive function is in part determined by health in infancy. Case and Paxson (2008) provide indirect evidence of the first relationship. Using longitudinal data from the UK, they show that the strong positive correlation between earnings and adult height\textsuperscript{21}, which is well established, falls greatly in magnitude and becomes insignificant once cognitive functioning in childhood, which is shown to be a strong predictor of wages, is controlled for. This is consistent with early life nutrition producing both cognition and height and only the former impacting on earnings. After dealing with bias arising from cognition measured in adolescence being a product of (endogenously chosen) schooling, Heckman et al (2006) find that cognitive functioning is an important determinant of schooling, employment, occupation and wages.

So, cognitive functioning in childhood has economic consequences. Do we know that cognition itself is contingent on health in infancy? Neuroscience has identified through animal experiments the biological and neurological processes that link undernutrition, as well as nutrient deficiency and exposure to toxins, in \textit{utero} and infancy to impeded development of the brain (Grantham-McGregor, Cheung et al. 2007). Birth weight, which is an indicator of exposure to health risks \textit{in utero}, particularly nutritional deprivation, is the most frequently used indicator of the health of humans at birth. Epidemiological studies confirm that low birth weight is associated with low IQ (Breslau, DelDotto et al. 1994), along with a host of other child health problems including asthma (Nepomnyaschy, Reichman 2006), behavioral problems including Attention Deficit Hyperactivity Disorder (ADHD) (Hayes, Sharif 2009, 2009).

\textsuperscript{21} In the US as well as the UK, someone who is one inch taller has, on average, 1.5-2\% higher earnings (Case and Paxson, 2008).
Loe, Lee et al. 2011), slower motor and social development (Hediger, Overpeck et al. 2002) and depression (Costello, Worthman et al. 2007). While consistent with the neuroscience, one cannot read too much into a simple correlation since low birth weight could reflect behaviour of the mother during pregnancy that is correlated with later investments in the child that influence cognitive functioning. Variation in birth weight within siblings and twins has been used to reduce the risk of such bias. This approach has produced evidence that the smaller sibling or twin tends to have lower IQ at age seven in Scotland (Lawlor, Clark et al. 2006) and on entrance to the military in Norway (Black, Devereux et al. 2007). The latter finding is more difficult to interpret since IQ in young adulthood could reflect differential investment in education.

In low and middle income countries, where the nutritional and micronutrient deficiencies are obviously much more pronounced, there is clearer evidence from randomly assigned nutrition supplementation programmes that better nutrition improves cognitive functioning, and also raises educational attainment (Grantham-McGregor, Powell et al. 1991, Pollitt, Gorman et al. 1993, Walker, Chang et al. 2005, Grantham-McGregor, Cheung et al. 2007). Currie (2009) cites evidence showing that even in the US children of mothers included (not randomly) in a nutritional programme during pregnancy achieve higher test scores.

Heckman and his collaborators present evidence demonstrating that noncognitive skills developed in childhood are as important, possibly even more important, than cognitive functioning in explaining economic outcomes in adulthood (Heckman, Stixrud et al. 2006, Heckman 2007, Cunha, Heckman 2009). Noncognitive capabilities refer to personality traits, such as self-esteem, perseverance, dependency, consistency, patience and optimism, which may be considered to be determinants or aspects of preferences over risk and the timing of consumption. Currie (2009) points out that some noncognitive skills are closely related to, or are highly contingent on, mental health conditions. She cites a number of studies presenting
evidence that child behavioural problems, such as ADHD and aggression, are strong predictors of lower cognitive functioning, educational attainment and economic outcomes. Most of the studies control only for observables, but Currie and Stabile (2006) use sibling fixed effects and find that children in both the US and Canada with high ADHD scores at younger ages had lower cognitive functioning (math and reading test scores) at age 11, and were more likely to be admitted to special education and to have repeated a grade. Given that behavioural mental health problems are so prevalent in children, they are potentially an important part of the link from child health to adult economic circumstances.

While the evidence base does need strengthening, we believe that there are sufficient conceptual grounds for expecting health in infancy and early childhood to emerge as an important constraint on the formation of cognitive and noncognitive skills that are increasingly recognized as important determinants of labour market success. Within the framework developed by Heckman and Cunha (2008), the importance of these skills in the generation and reproduction of economic inequality derives from the fact that they are malleable. Skills are produced through parental investments, the level and efficiency of which are likely to depend on the socioeconomic environment of the child’s family and neighbourhood. Heckman (2007) hypothesizes that there may be complementarities between investments in health and cognition. Bad luck of the draw which leaves a child with a deficit of one may make it more difficult to raise the other through investments. An early life health shock could leave a child frail, with limited capacity to respond to the stimuli that can raise cognitive functioning. Facing a higher price for a marginal gain in functioning, parents may invest less in development of the child’s skills. Sickly kids may be doubly penalized with a deficit in both health and human capital.
3.4.2 Education

Education is a potential conduit that links income to health not only over the life cycle but also across generations. Health problems in childhood may directly constrain the acquisition of education, as well as weaken incentives for investment in schooling, with long run consequences for income. If the children of poorer and less healthy parents are more likely to experience illness, then the interference of health capital in the acquisition of other forms of human capital could contribute to the intergenerational transmission of income (Currie, 2009).

Health determination of the income distribution may operate with a very long lag.

According to life course epidemiology (Kuh, Wadsworth 1993, Kuh, Ben-Shlomo 1997, Wadsworth, Kuh 1997), childhood illness, which may arise from social deprivation, not only has a permanent effect on health, it also interferes with education. As a result, occupational opportunities are limited in young adulthood, which may further reduce adult health, and lifetime earnings potential is constrained. Health and income in adulthood are correlated because they are both determined by childhood illness.

The most straightforward way in which health may impact on education is through the interruption of schooling. In low income countries, this could be an important constraint. In high income countries, it seems less relevant. Currie (2009), citing Grossman and Kaestner (1997), notes that differences in school absence between healthy and unhealthy US kids is too small for this to lead to a strong correlation between health and educational attainment. Any impact of child ill-health on knowledge and skills acquisition is more likely to operate through the channel examined in the last sub-section—constrained cognitive functioning and impaired efficiency of learning.

Using a cohort of Britons born in 1958 and controlling for childhood SES, Case et al (2005) find that a health problem in childhood is correlated with lower educational attainment. But
poor health and low education could both result from parents who make little investment in the human capital of their offspring. A partial solution is to go one stage back in the child’s development when parental behaviour can exercise less influence and examine the association between birth weight and educational attainment. We noted above the evidence of an impact of birth weight on cognition. If this is the channel, we now need to establish the next link in the chain, to educational outcomes. There are many epidemiological and social science studies showing that children with very low, or even low, birth weight tend to perform poorly in school (Saigal, Szatmari et al. 1991, Hille, Ouden et al. 1994, Currie, Hyson 1999, Case, Fertig et al. 2005, Kirkegaard, Obel et al. 2006). That the relationship is causal is given credence by sibling and twin differences studies of sufficiently large samples from high income countries across three continents that show lower educational attainment of the smaller child at birth (Behrman, Rosenzweig 2004, Black, Devereux et al. 2007, Oreopoulos, Stabile et al. 2008, Royer 2009, Lin, Liu 2009, Johnson, Schoeni 2011). There is, however, variation in the magnitudes of the estimates. For example, using US PSID data, Johnson & Schoeni (2011) find that low birth weight increases the probability of dropping out of high school by one-third, with part of the effect appearing to operate through impaired cognitive functioning. Royer (2009), using data on twins who both become mothers in California, finds a rather small average effect.22 A claimed feasible 250 gram increase in birth weight would raise schooling by only 0.04 of a year.

A couple of US studies only find significant or substantial effects for children born into poor families or neighbourhoods (Conley, Bennett 2001, Currie, Moretti 2007). This is consistent with wealthier parents being able to compensate for a health disadvantage in early life, while poorer parents lack the means to invest in medical care, or other health and educational inputs. Other studies conducted with British (Currie, Hyson 1999), Norwegian (Black, Devereux et

---

22 A priori, restriction of the sample to twin pairs that both become mothers risks selection bias although analysis conducted by the author suggests that this is not large.
al. 2007) and Canadian (Oreopoulos, Stabile et al. 2008) data find no evidence of this heterogeneity. Hasty attribution of this discrepancy to the equalizing effect of universal health care coverage in Europe and Canada that is absent in the US would be foolhardy. The 1958 British cohort would have access to only rudimentary medical interventions for the treatment of low weight babies by today’s standards (Almond, Currie 2012). But it could be that little treatment was available for rich and poor low birth weight babies alike in fifties Britain. The pertinent question is whether the gradient observed in the more recent US data arises from differential access to effective medical care, and possibly other corrective interventions. It may also be that the incentives parents have to make investments that compensate for poor child health differ between the rich and the poor. A particularly tight budget constraint may not stretch to investing in all offspring. It can then be optimal to concentrate investments on the child with the best chance at the expense of the child the parents observe to be frail and likely to struggle in life in any case (Almond, Currie 2012).

There is emerging evidence of a link between measures of school performance, or educational attainment, and in utero exposure to health risks through disease (Almond 2006, Kelly 2011), radiation (Almond, Edlund et al. 2009) and maternal alcohol consumption (Nilsson 2009). This is arguably a more convincing strategy to identify the effect of early life health on education than that of twin differences since it uses variation in infant health risks that is external to the family environment and so is more plausibly exogenous to factors that also impact on schooling. Even more convincing is evidence that treatment of children for disease raises school attendance (Miguel, Kremer 2004, Bleakley 2007).

---

23 In general, the investment response will depend on the technology of human capital production and parental aversion to inequality across children (Almond and Currie, 2011). When technology is such that early life health shocks can be offset by later childhood investments only to a limited extent, it can be optimal to reinforce the shock by withdrawing investment in the frail child irrespective of the degree of inequality aversion. This is more likely the lower is household income. See Appendix C and footnote 7 in Almond and Currie (2011).
The evidence is convincing that health status at the time of birth impacts positively on educational attainment. How does this effect operate? Does fetal distress permanently damage cognitive functioning, which interferes with knowledge acquisition and skills development throughout school? Or do children who are frail at birth subsequently develop health conditions and illnesses at pre-school ages that delay development and place the child at a disadvantage throughout its schooling? Or do health problems at pre-school and early school years lead to additional health problems in later school life at ages when qualifications must be obtained? Currie, Stabile et al (2010) are able to address these type of questions by using rich Canadian data following children from birth to young adulthood. Controlling sibling fixed effects, birth weight and, congenital and perinatal abnormalities, which always have effects, major physical health conditions and injuries in pre-school (0-3) and early school (4-8) years impact on educational attainment only because they raise the probability of experiencing similar conditions in later school years. The process is different with respect to mental health conditions. Behavioural problems at all ages, including pre-school years, directly reduce educational attainment. From this and other evidence Almond and Currie (2012) conclude that there is an important effect of early life health on human capital acquisition and later life outcomes. Exposure to adverse conditions in utero appear to exert a stronger effect than postnatal health, although mental health conditions identified pre-school have lasting effects.

Besides a technological effect of childhood ill-health on the acquisition of education, if the reduction in health is permanent, then there may be an incentive effect operating through life expectancy. A longer life may raise education by lengthening the period over which the return on this investment can be reaped (Ben-Porath 1967). This further strengthens the extent to which dispersion in health implies dispersion in earnings potential. There is an inequality increasing concentration of health and human capital in the same individuals.
3.4.3 Fetal origins hypothesis

Early life health potentially determines later life economic outcomes not only via education and skills acquisition but more directly through health problems that interfere with productive work in adulthood. The *fetal origins hypothesis*, proposed by David Barker, is that chronic diseases, principally coronary heart diseases but also related diseases such as type 2 diabetes, originate in nutritional deprivation in gestation and infancy (Barker 1992, Barker, Godfrey et al. 1993, Barker 1995). Nutrition-induced stresses placed on the fetus at critical stages of development alter the physiology of vital organs, particularly the heart, which makes them susceptible to failure in middle and old age inducing the onset of chronic disease.\(^{24}\) Metabolism can also be detrimentally affected by *in utero* nutritional deprivation such that the risk of obesity is raised.

One cause of fetal stress that has been exploited to estimate long term economic effects is *in utero* exposure to the 1918-19 Spanish flu (Almond 2006), which has been shown to increase the incidence of stroke, diabetes, hearing, seeing and mobility impairments (Almond, Mazumder 2005), as well as cardiovascular disease (Mazumder, Almond et al. 2010). As is clear from the evidence reviewed in section 3.3.5, chronic illnesses such as these have negative impacts on employment and earnings. Economic consequences of the *fetal origins hypothesis* has recently (2013) received much attention (Currie 2009, Almond, Currie 2011, Almond, Currie 2012). We discuss some of the evidence in the next sub-section.

3.4.4 Economic consequences of early life ill-health

The previous three sub-sections establish that ill-health in early life and childhood constrains the acquisition of non-health human capital and directly impinges on adult health. With these disadvantages, one expects adults who were sick, or exposed to health risks, as infants and

---

\(^{24}\) The theory is not free of sceptics in medicine and epidemiology. See, for example, Paneth and Susser (1995) who call for clearer elucidation of precise physiological mechanisms and their testing.
children to be less well-off. What is the evidence that frailty of health at the beginning of life leads to economic disadvantage later in life?

Analyses of the 1958 British birth cohort establish that low birth weight is not only associated with lower educational attainment but also with a lower employment, a greater likelihood of being engaged in manual labour and lower wages (Currie, Hyson 1999, Case, Fertig et al. 2005). Sibling and twin fixed effects studies produce evidence that lower birth weight reduces wages in Minnesota (Behrman, Rosenzweig 2004), reduces employment and earnings in the US (Johnson, Schoeni 2011)\(^\text{25}\), reduces earnings (but not employment) in Norway (Black, Devereux et al. 2007)\(^\text{26}\), increases welfare dependency in Canada (Oreopoulos, Stabile et al. 2008, Currie, Stabile et al. 2010), and may (Currie, Moretti 2007) or may not (Royer 2009) increase the likelihood of living in a poor neighbourhood of California.

Johnson and Schoeni (2011) find that only around 10% of the effect of low birth weight on earnings operates through years of schooling.\(^\text{27}\) That most of the health effect does not operate via human capital is consistent with the findings of other studies (Persico, Postlewaite et al. 2004, Luo, Waite 2005, Smith 2009). Using the 1970 British Cohort Study, Conti, Heckman and Urzua (2010, 2010) do not find any selection into post-compulsory schooling on the basis of child health for males and only a weak effect for females, while there is very strong selection on cognitive and noncognitive skills. There is a direct effect of child health on wages for males, while the weak wage effect for females does run via education.

---

\(^\text{25}\) Low birth weight is estimated to reduce labour force participation by five percentage points and earnings (given employment) by around 15%.

\(^\text{26}\) A 10% increase in birth weight is estimated to raise earnings of the full-time employed by 1%, an effect equivalent to that of about 3 months of education.

\(^\text{27}\) In interpreting this one should bear in mind that birth weight is estimated to have a small and insignificant impact on years of education. It is on the probability of high school graduation and on test scores that it impacts strongly. If these intermediate outcomes were considered, then presumably the proportion of the effect of birth weight on earnings that operates through human capital would be revealed to be greater.
The health measures employed by Conti et al are height and head circumference at age 10. While these should pick up nutritional deprivation at early ages, they will miss many other child health conditions. Estimating the long term economic impact of child health more generally defined is made difficult by the scarcity of suitable longitudinal data and the challenge of measuring the general health status of children. Using the US PSID, Smith (2009) estimates that an adult retrospectively reporting excellent or very good health in childhood earns 24% more than his or her sibling who reports less than good health in childhood. There are no differences in education, so again most of the effect, if we are prepared to label the earnings difference as such, does not appear to operate through human capital.

In an intriguing paper, Almond (2006) finds dramatic effects of in utero exposure to the 1919 Spanish flu on economic outcomes. Earnings of male workers were reduced by 5-9%. Applying standard estimates of the return to education to the impact on schooling (length reduced by five months, on average, and high school graduation probability down by 13-15%), the indirect effect through education explains around one half of the earnings effect. So, while there is support for the life course model that proposes that the infant health effect operates through education, this does not tell the full story. There appears to be a substantial direct effect from fetal conditions to earnings.

Welfare payments, which include disability insurance, are higher for those exposed to the flu. There is also evidence of a substantial effect through occupation. For men, exposure results in a drop of around 6% on a hierarchical index of occupational status. Total income is reduced by about 6.4% for men and the probability of being poor is increased by as much as 15
percentage points. Identification of these effects from pre-birth conditions cuts through the endogeneity problems that plague the evidence on the economic consequences of ill-health reviewed in section 3.3.5. The limitation is that the link between health and economic outcomes is less transparent. The identification requires a leap of faith that it is the Spanish flu and not some other peculiarity of the 1919 birth cohort that is responsible for the effects.

Subject to this caveat, this study has important implications for interpretation of the distribution of income and its association with health. Sizable differences in incomes, and in its education and occupation determinants, appear to be attributable to differences in health conditions at the very beginning of life. Policies that can improve the early life conditions of the most vulnerable infants can potentially compress the distribution of income decades later by ensuring that there are fewer physically frail individuals who can become the most economically deprived. Almond (2006) draws attention to the tremendous racial disparity in early life conditions in the US, where nonwhites are exposed to twice the infant mortality rate of whites. His estimates suggest that this health inequality may not only be a consequence of current economic inequality but a potentially important cause of future economic inequality.

Overall, the evidence points to the health-related environment into which a child is conceived and delivered exerting an effect on lifelong economic opportunities. As Currie (2009) notes, it is difficult to gauge the magnitude of the effect given the variety of sources of variation in health risks and measures that have been employed by researchers and the differing study contexts. The evidence is sufficient to conclude that it is not a negligible effect. A further

---

28 Poverty is defined as an income below 150% of the poverty line. The effects presented here for occupation and income are computed as the average of the effects for the three census years presented in Table 1 each scaled by a one-third infection rate and expressed relative to the control group means in Table 2 in Almond (2006).

29 Additional evidence of long term economic effects of early life health conditions includes Nilsson’s (2009) finding that increased in utero exposure to alcohol due to mothers drinking more following a liberalization of licensing laws in Sweden not only reduced years of schooling but also decreased earnings and increased welfare dependency. From the effect of a programme to eradicate hookworm in the American South in the first quarter of the twentieth century, Bleakley (2007) estimates that infection with the disease in childhood reduced the probability of school enrollment by 20% and reduced wages by around 40%.
tentative conclusion is that the effect through adult health is stronger than that through human capital. This is sometimes claimed as support for the *fetal origins hypothesis*. Strictly, it is not. The hypothesis claims a causal link from fetal health to the onset of chronic disease in middle age. A number of studies examine the health of reasonably young adults and the measures used often do not identify the cardiovascular conditions that are the main diseases triggered by fetal stress according to the hypothesis. Adult health can be related to health conditions in childhood also because health is persistent, which is a feature of the *life course model* of epidemiology. Health in early life is relevant to economic circumstances in adult life. Establishing precisely why and to what extent presents a challenging research agenda to both economists and epidemiologists.

### 3.5 Health and occupation

The *life course model* identifies entry to the labour market as an important stage at which the relationship between health and socioeconomic status is strengthened. Frail young adults face a narrower choice of entry level jobs both because their education has been constrained by ill-health and because persistent health conditions directly impede productivity, or provoke discrimination. The evidence reviewed in the previous section suggests that the direct productivity effect is stronger than the education effect. The income gradient in health may then partly reflect selection of the less healthy into lower paid jobs, as well as sickness impeding movement up the career ladder.

This mechanism has been the focus of the epidemiological literature that has considered the extent to which causality runs from health to socioeconomic status (SES). The literature generally concludes that health-related occupation selection and social mobility are of insufficient magnitudes to make important contributions to the observed socioeconomic gradient in health (Power, Matthews et al. 1996, Power, Matthews et al. 1998, Chandola, Bartley et al. 2003). In this literature, SES is typically measured by occupation, social class or
employment grade. What is found is that while job changes are related to health, movements of the more healthy into ‘better jobs’ are insufficient to explain the observed health disparity across the occupational hierarchy (Chandola, Bartley et al. 2003). The same need not follow with respect to explanation of health differences across more economic dimensions of SES, such as earnings or income (Adda, Chandola et al. 2003). As hopefully is clear by now, ill-health can impact on income through many channels other than occupation and so even if this were not an important mechanism it certainly does not follow that the income (rather than the occupation) gradient in health is mainly attributable to causality from income to health, rather than vice versa. Researchers from the economics discipline are generally more sympathetic to the selection hypothesis (Deaton 2002, Smith 2004, Deaton forthcoming). In particular, the potential for the evolution of both an individual’s career and health to be determined by his early life health and other experiences, as discussed in the previous section, is recognized and emphasized (Case, Paxson 2011).

Health-related selection of occupation need not necessarily steepen the income gradient in health. If productivity varies with health and productivity is unobservable, then wages will vary with health only if the latter is observable (Strauss, Thomas 1998). If health is not observable, then the healthy would be expected to sort into occupations in which productivity is less difficult to observe. Particularly in low income countries, the healthy may stick to self-employment in which there is no issue of productivity verification and earnings are not set at some average over higher and lower levels of productivity. But such occupations may offer lower earnings potential; perhaps because there is less capital per unit of labour. Wage differences by health would be more compressed than in a situation in which there was no sorting.

The evidence reviewed in section 3.2.4 generally provides estimates of the impact of health on the wage rate conditional on occupation. But there could be an additional effect on
earnings through occupation itself. The evidence on job change after the onset of a health condition is relatively sparse. Now rather old data reveal that around one quarter of males and one fifth of females in the US with a work-limiting health problem report having moved to jobs more compatible with their condition (Daly, Bound 1996). Older workers and high school dropouts are less likely to change jobs. Charles (2003) argues that older workers are less likely to retrain to accommodate a disability since they have less time remaining to reap a return on this investment. He finds evidence consistent with this hypothesis from the US PSID. The low educated may be less likely to adapt because they lack the general human capital that raises the efficiency of specific investments. It is more difficult to move from manual to non-manual labour if basic reading and writing skills are lacking. This may partly explain why lower educated, lower skilled and lower paid workers are more likely to exit the labour force when struck down by a disability (Banks 2006, Halla, Zweimuller 2013, García Gómez, van Kippersluis et al. 2013) and, in the US, are less likely to recover earnings lost after initial onset of disability (Charles 2003).

3.6 Health and household income

The impact of health on the distribution of household incomes may differ from that on the distribution of individual incomes for two principal reasons. First, health may affect the formation and dissolution of households. Second, illness of one household member may provoke a response from the labour supply of others.

Poor health may make it more difficult to find a partner. Limitations in functioning, caring needs and reduced expected longevity may make a disabled or chronically ill person a less attractive proposition in what some economists refer to as the marriage market. On top of the direct effect, there may be a reinforcing indirect effect operating through the impact of health on human capital accumulation and earnings potential; marrying a sicker person, on average, means marrying a poorer person. The marriage vows of ‘in sickness and in health’ appear to
recognize the threat that illness poses to marriage. If less healthy people do have fewer opportunities to find and keep a partner, then one would expect that there will be health-related sorting. Then lower than average earnings of a disabled or sickly person will be compounded, not compensated, by the earnings of his or her spouse.

While the idea that health influences marriage prospects has been around for some time (Sheps 1961, Carter, Glick 1976), there is little convincing evidence with which to judge its empirical validity. In part, this is because of difficulty of separating healthy selection into marriage from the potentially beneficial effect of marriage on health (Goldman 1993). Longitudinal data are required. Fu and Goldman (1994) find little evidence that health predicts the marriage behaviour of young American adults. For US women there is evidence of health selection. Among young women not in full-time employment, those in better health are more likely to marry and less likely to break up (Waldron, Hughes et al. 1996) and in a Californian sample of siblings the sister with the lower birth weight is three percentage points less likely to be married when she gives birth (Currie, Moretti 2007).

Recognition of potential health gains from marriage introduces the possibility that there is negative health selection into the institution. The less healthy have more to gain from marriage. Lillard and Panis (1996) find evidence of such adverse selection among US men – the less healthy (re)marry sooner and remain married for longer. However, there is also selection on unobservables correlated with good health and this dominates, such that married men are healthier than their unmarried counterparts.

Within a household, ill-health of one partner could provoke two conflicting motivations for labour market activity of the partner. On the one hand, reduced earnings of the disabled partner will generate an income effect that will motivate the spouse to replace those lost earnings through increased work effort. This is the added worker effect that is familiar in the
unemployment literature. On the other hand, the disabled person’s productivity may be reduced not only in the labour market, but also within the household. Limitations in functioning may reduce capacity to wash, dress and feed oneself. Meeting these caring needs will place demands on the spouse’s time. A priori, one cannot say which effect will dominate.

There appears to be a gender difference in the relative magnitude of the two effects but the direction of this bias is not always consistent. Most US evidence finds that women are more likely to participate in employment when their husbands fall ill but, if anything, male spouses are less likely to participate (Berger 1983, Berger, Fleisher 1984, Charles 1999, Van Houtven, Coe 2010). But Coile (2004) finds no effect on the female spouse and only a small increase in employment of men whose wives fall ill. The employment response of the spouse has been found to depend on the type of health condition and the initial labour supply of the spouse (Blau, Riphahn 1999, Siegel 2006).

There is evidence from Germany (Riphan 1999), Spain (García Gómez, Lopez-Nicolas 2006) and the Netherlands (García Gómez, van Kippersluis et al. 2013) of ill-health reducing household income by more than the fall in the personal income of the person experiencing the health shock. For example, the Dutch study finds that an acute hospital admission reduces household income by 50% more than the reduction in income of the person admitted to hospital.

3.7 Health and wealth

The impact of health on economic inequality may go beyond the wage and income distributions to the wealth distribution.30 If ill-health reduces income through one or more of the mechanisms identified in the previous sub-sections, then opportunities to accumulate

30 Identification of the distribution of wealth itself can require taking account of the distribution of health, or rather mortality. When one only observes wealth of the deceased, in the form of inheritances, then mortality multipliers need to be applied in order to infer the distribution of wealth among the living. Differential mortality may be taken into account (Atkinson, Harrison 1978)
wealth over a lifetime will be constrained. Because the effect is accumulated, permanent differences in health will create greater variance in wealth than in income. In addition, ill-health may force depletion of wealth to pay for medical or nursing care. Less obviously, health may affect wealth through life expectancy and consequent saving incentives. The horizon and uncertainty effects of increased longevity both raise saving, while they have contradictory effects on labour supply (see section 3.3.4). Those expecting to live for longer will accumulate more wealth both to provide for an extended old age and because they face a lower risk of dying before having the opportunity to enjoy their savings.

Consistent with these mechanisms, the relationship between health and wealth is particularly strong. PSID data reveal that the median wealth in 1994 of a household whose head was in excellent health ten years earlier was 268% greater than the median wealth of a household whose head had been in poor health (Smith 1999). This wealth inequality grew both with the lapse of time since the difference in health was recorded and with age, consistent with differential rates of wealth accumulation by health. Whether the strong positive relationship between health and wealth is due to causality from health to wealth, from wealth to health or arises from spurious correlation has been examined in a number of US studies.

The much-cited paper by Adams et al (2003) analyzes panel data on a sample of the US population aged 70+ (Asset and Health Dynamics among the Oldest Old – AHEAD). Focus on the elderly eliminates differential earnings as a mechanism through which health may contribute to differences in wealth accumulation. The null of no causation from both current and previous health – indicated by 19 conditions and SAH – to the change in wealth is rejected.31 Their rigorous analysis leads the authors to emphasize that, while the result is consistent with a causal effect from health to wealth, they cannot rule out the possibility that it

31 The null is rejected for total, liquid and non-liquid wealth for couples in which both partners survive and single households (except non-liquid wealth). The null is not rejected in a number of cases for couples experiencing a death (see Adams et al (2003), Table 11)
reflects model misspecification and/or time invariant unobservable factors driving the evolution of both health and wealth.

Michaud and van Soest (2008) overcome both limitations and provide even more conclusive evidence of a causal effect from health to wealth. They use the HRS and allow for causality operating contemporaneously and with lags in both directions, along with unobservable heterogeneity. Health is measured by an index constructed from principal components analysis of SAH, major and minor conditions, Activities of Daily Living, depression score and body mass index. Health of both the husband and the wife are found to impact on household wealth. The effect of the wife’s health is immediate, while that of the husband’s health is delayed. This is due to a gender difference in the type of ill-health that impacts on wealth. For both sexes there is a delayed impact of physical ill-health. But only the mental health of females has an impact on wealth and this is immediate. Evidence of a causal impact of health on wealth is stronger for households that lack health insurance coverage, particularly those in which the wife succumbs to mental ill-health, suggesting that depletion of assets to pay for medical care is an important part of the effect.

While the Michaud and van Soest study gives good reason to believe that there is an effect of health on wealth, at least in the older US population, its use of a health index has the disadvantage of not producing an estimate of easily interpretable magnitude. Without allowing for unobservables that condition the evolution of both health and wealth, Smith (2004) estimates, using HRS data, that the income loss, medical expenses and consequent forgone of interest arising from the onset of a major health condition (see section 3.3.5) accumulate over eight years to an average loss of wealth of almost $50,000. Most of this is

---

32 Instruments are required in models that allow contemporaneous effects. The onset of major health conditions (cancer, heart condition, lung disease and stroke) is used under the assumption that these critical illnesses impact on wealth only through health and are not contemporaneously affected by changes in wealth. Essentially, this is the same identification assumption used by Smith (2004) and Wu (2003).
due to reduced earnings. Consequently, the wealth loss is considerably lower ($11,350) for the older AHEAD cohort analyzed by Adams et al. The wealth loss from the onset of a minor health condition is also much smaller ($11,500). Recognizing that one fifth of Americans aged 50+ experience the onset of a major health condition over an eight year period and a further 30% incur a minor condition, Smith argues that the consequent wealth losses represent substantial effects of health on the distribution of wealth. Consistent with Michaud and van Soest (2008), the magnitude of the effect that does not operate though earnings losses is larger when it is the wife than experiences the illness (Wu 2003).33 This is explained by assets being run down to pay for general living expenses when the wife is no longer fit to perform household chores. This effect is not observed when the husband’s health deteriorates.

Overall, the evidence is convincing that health constrains the accumulation of wealth and illness speeds its depletion. The magnitude of the effect is likely to differ with the nature of the health condition and the means of financing both pensions and medical care.34 In the US, for which most evidence is available, the effect seems substantial. Variation in both health levels and rates of health depreciation with age may make substantial contributions to inter-household inequality in wealth holdings.

### 3.8 Summary

Understanding the effects of health on income and wealth is important in relation to both explanation of the distributions of income and wealth, and interpretation of the economic gradient in health. We have identified a number of pathways through which health potentially

---

33 This evidence comes from analysis of only the first two waves of the HRS and the contemporaneous relationship between changes in health and changes in wealth.

34 Hurd and Kapteyn (2003) propose and confirm that the relationship between wealth and health at older ages is weaker in a country such as the Netherlands, where a greater share of retirement income is obtained from annuities, than in a country such as the US, where savings and assets are more important sources of financing consumption in old age. Differential financing of medical care, and not only the source of retirement incomes, may contribute to this result. One expects a stronger correlation between wealth and health in countries, such as the US, with less comprehensive public health insurance and where, until the 2010 Affordable Care Act, private insurance premiums could be related to pre-existing conditions.
impacts on income and wealth, and the evidence suggests that many of these are empirically important. Ill-health can lead to a fall in wages at the margin, but wages more likely drop through job changes. Institutional constraints on wage flexibility in high-income economies, including anti-discrimination legislation, can limit the wage response but at the cost of increasing the employment loss. Disability insurance further increases the employment effect of ill-health through sizable incentive effects. Because of a spillover effect on the labour supply of the spouse, the impact of ill-health on household income can be substantially larger than that on the earnings of the disabled person. This, combined with illness possibly reducing the likelihood of forming and maintaining marriage partnerships, although the evidence on this is mixed, means that the contribution of ill-health to inequality in household incomes need not be less, and may even be greater, than its contribution to inequality in personal incomes. Perhaps the most important conclusion emerging from the literature, and emphasized by others (Heckman 2007, Currie 2009, Almond, Currie 2012), is that ill-health can have a very long reach from exposure to health risks in childhood to constrained economic opportunities in adulthood. By constraining human capital acquisition through education and skills formation, but also through persistent and delayed effects on health in adulthood, early life and childhood health experiences can be important determinants both of the distribution of income and the observed income gradient in adult health.

There is a good deal of evidence relevant to most of the mechanisms through which health may impact on income and wealth but inferring the magnitude of each effect is far from easy. Not least, this is because, even with respect to a particular mechanism, there is not a single effect but many. Health is not unidimensional. Different dimensions of health will impact on income and wealth through different routes and to different degrees. Using a general measure of health, such as self-assessed health, may provide some average effect over different types of health problems but the usefulness of this average is questionable. Related to this,
measurement error in health variables has been a substantial obstacle to obtaining credible estimates of health on labour market outcomes. But this problem is receding. Longitudinal surveys, such as the HRS and its equivalents, increasingly contain detailed measures of specific health conditions allowing researchers to exploit the timing of illness onset to identify the economic consequences of intrinsically interesting changes in health with precise medical meanings. Also promising is the increasing access of researchers to linked administrative registers on hospital admissions, social insurance and tax files that drastically reduce measurement error and provide very large samples from which the effects of specific medical conditions can be identified.

Ill-health reduces income and wealth. The contribution of health to economic inequality depends upon how it is distributed. If health variation were random, then it would add to the dispersion of income (wealth). But in this case the additional economic inequality may not be considered socially objectionable. Losing income as a consequence of illness may be seen as unlucky but not unjust. Insurance would be called for on grounds of efficiency. But the consequences for economic inequality – both the nature of the impact and its normative interpretation – would be quite different if ill-health is not distributed by the roll of the dice. The next section examines whether income and wealth exert causal effects on health. Irrespective of whether such effects exist, if individuals with lower potential incomes are more likely to fall sick, then the income distribution will be skewed even more to the disadvantage of the poor. For example, assume that low education both reduces income and increases the likelihood of sickness. The poor are more likely to get sick and because they are sick they become even poorer. The income distribution gets stretched by the unequal incidence of illness and its impact on income. The economic inequality generated by biases that place the socially underprivileged at a health disadvantaged and then impose an economic
penalty for this is more likely to be considered morally objectionable than inequality arising from even handed luck in the distribution of illness.

In addition to the uneven incidence of illness, the effect of health on economic inequality is likely exacerbated by heterogeneity in the impact of health on income across the income distribution. The employment and earnings of low skilled, low paid workers is more contingent on their health than is the case for higher paid professionals. Not only do the socially disadvantaged face a higher incidence of illness, but they are more economically vulnerable to it. Disability insurance provides a safety net but its disincentive effects are stronger for the low paid and they are more responsive to these incentives since, at least in recent decades, labour market opportunities are deteriorating for this group. As a result, the loss of employment following ill-health is not only a contributor to economic inequality but a consequence of it.

We conjecture that ill-health contributes to economic inequality not merely by adding noise to the distribution of income (wealth) but by further reducing the incomes (wealth) of those that would be located toward the bottom of the distribution in any case. This could occur even without low income or wealth reducing health. We now turn to the question of whether there is a causal effect from the economic to the health domain of well-being.

4. Economic determination of health inequality

4.1 Overview

If, as one would expect, health is a normal good, then the financially better-off will demand more of it. Whether this inflated demand is realised will depend on how health enhancing and depleting goods are allocated. If medical care is delivered through the market, then the rich will be both willing and able to afford more effective treatment when illness strikes. But few countries, particularly high income ones, leave the distribution of health care entirely to the
market. Public provision of care to the poor and elderly, or even universally to the entire population, should constrain health differences that arise from variation in the willingness and ability to pay for medicine. But other goods that are beneficial to health, such as quality housing, safe neighbourhoods and education, are at least partially allocated by the market and provide scope for income to ‘buy’ health. The direction of the relationship is not, however, unambiguous. Whether inequality in health reflects economic inequality will depend on the extent to which the greater demand of richer individuals for health is offset by their higher demand for the pleasures of alcohol, smoking and rich foods that higher income makes affordable.

Whether the better-off should be expected to be in better health will also depend on the source of their economic advantage. If it arises from higher earnings potential, then the health effect is ambiguous. This insight emerges from Grossman’s (1972a) health capital model according to which health is demanded for a direct utility benefit – feeling sick is uncomfortable – and a production benefit – less time is lost to sickness and so more is available for work. Health is produced by investment in medical care, exercise, healthy eating, etc. In the pure investment version of the model, which incorporates the production benefit only, a wage increase has two conflicting effects on health. A higher wage implies a higher value of a given increase in productive time, which would lead higher waged individuals to invest more in health. But the marginal cost of the time input into health investment also increases. More earnings are lost visiting the doctor, jogging, etc. The net effect is positive provided market goods, e.g. medical care, nutritious food, are used in the production of health in addition to the individual’s own time input (Grossman 2000). But in the pure consumption version of the model, which confines attention to the direct utility benefit of health, it is not sufficient that the time cost of producing health is less than the total cost in order that the substitution effect of a wage increase is positive. The relative intensity with which time is used in the production of health
must be less than the relative time input into the production of other commodities that generate utility. Otherwise, a wage increase implies a rise in the relative price of health. Maintaining one’s health is likely to be more time intensive than many other activities generating sources of utility and so a negative pure wage effect certainly cannot be ruled out.

Answering the empirical question of the extent to which economic advantage bestows health advantage is complicated greatly by the multitude of mechanisms identified in the previous section through which health impacts on economic circumstances, as well as a plethora of unobservables, such as risk attitudes, time preferences and genetics, that can influence investments in health and in other human and financial capital. Fixed effects methods deal with the latter problem but are powerless against the former simultaneity problem. Early attempts to tackle both problems tended to rely on instruments of questionable validity. More recently, researchers have confined attention to the health response to more plausibly exogenous sources of variation in income or wealth, such as that arising from sudden policy reforms, stock market volatility or windfall gains. The weakness of this strategy is that the economic variation used, while exogenous, does not correspond to that which can plausibly be expected to impact on health. The problem is exacerbated by the fact that health does not respond immediately to a change in demand. Even if a windfall gain from a stock market boom, inheritance or lottery win does induce a rise in the desired level of health, it will take years of investment in preventive medical care, diet, etc. to realize this. Even long panels may be insufficient to observe this process. There is a greater chance of identifying the income effect on health determinants than on health itself.

Many chronic health problems are unlikely to be provoked by sudden changes in income or wealth but may (partly) result from long term exposure to unhealthy living conditions.

---

35For example, Ettner (1996) instruments the wage rate with work experience and the state unemployment rate, and unearned income with parental and spousal education. Work experience may be correlated with the evolution of health and parental education could have a direct effect on health.
experienced by the poor. The lags involved in such an effect, along with the obvious endogeneity issues, make its identification challenging. Most of the evidence is not informative of such long term relationships and this needs to be kept in mind in interpreting the evidence from shorter term variation that tends to show no, or a weak, effect of income or wealth on health, at least in higher income countries.

We begin our review of the evidence in this section by examining the impact of income and wealth on health in adulthood. Most of this evidence comes from high income countries. We then turn to mechanisms and look at the evidence that economic resources impact on health behaviour and utilization of medical care. The penultimate subsection examines evidence for an impact of household economic circumstances in child health. Much of this evidence is from lower and middle income countries, or refers to low income populations in high income countries.36

4.2 Income and wealth effects on adult health

4.2.1 Tests of no causal effects

Perhaps the most influential examination of the economic determinants of health conducted in recent years is the study by Adams, Hurd et al (2003). Recognizing the difficulty of finding plausibly exogenous instruments for economic circumstances that provide variation relevant to mechanisms through which a causal effect may operate, the authors concentrate on the less demanding task of testing for the absence of causal effects of income and wealth (and other dimensions of SES) on health among elderly (70+) Americans (see Table 18.5 for details of this study and other evidence relating to the US). Focus on an elderly sample neatly sidesteps the reverse causality from health to earnings that would be likely in a sample of working age, and this is further avoided by adopting the concept of Granger (1969) causality and testing

36 We purposefully do not cover macro studies that attempt to identify the impact of country GDP on the mortality rate, e.g. (Pritchett, Summers 1996), since these tells us nothing about whether income is a determinant of health within any country.
whether, conditional on lagged health, current health is uncorrelated with lagged income (wealth). This involves imposition of an assumption that there is no contemporaneous impact of income or wealth on health, which the authors argue is plausible and strive to weaken by conditioning on a battery of health conditions that are presumed, in a medical sense, to be causally upstream of the condition for which causality is being tested.

The null that lagged income (wealth) does not predict health is not rejected for most conditions, including acute, sudden-onset conditions and mortality. This is interpreted by the authors as consistent with there being no causal effect of income and wealth on most health outcomes. The hypothesis of no causal effect of wealth is rejected for the incidence of mental health problems. The results for chronic and degenerative diseases are mixed. The authors argue that since treatment of mental and chronic illnesses are often not (fully) covered by Medicare, ability to pay may be a causal factor in the determination of these conditions.

Stowasser et al (2012) revisit the analysis by applying the same Granger causality tests to the original data source extended to a longer observation period, younger cohorts at a given age and younger ages (50+ rather than 70+). The last extension generates some variation in health insurance status that was not present in the older, Medicare-eligible sample included in the original study. With these changes, the null that health is conditionally independent of lagged income and wealth is rejected for a much larger number of conditions, leaving only a minority of conditions for which it is not rejected. This result is problematic for the approach since rejection of the null can arise either because there is indeed a true causal effect, or simply because of common correlation with omitted unobservables. So, while the original study tended toward the conclusion of no causal effect of income or wealth on health, analysis of

---

37 Since correlation through omitted common determinants is not ruled out, this interpretation of the test outcome is based on the presumption that there is no offsetting bias from unobserved heterogeneity that confounds a true causal effect sufficiently such that the net association is insignificantly different from zero (Heckman 2003, Stowasser, Heiss et al. 2012).

38 Adda et al (2003) question the plausibility of this interpretation because they find similar results using the same tests applied to Swedish and UK cohorts that are fully covered.
more data leaves one in the unfortunate situation of being unable to conclude either that there is no causal effect or that there is a causal effect.

[Table 18.5 here]

### 4.2.2 Estimates of effects
Panel data methods can be used to deal with time invariant unobservable determinants of health and income that cloud the conclusions that can be drawn from Granger causality analysis, although this is not straightforward using the nonlinear estimators appropriate for modeling often discrete and categorical health measures and when dynamics and long term relationships are taken into account. Using British data to estimate a dynamic random effects model of SAH, Contoyannis et al (2004) find that health varies with income averaged over time but not with current income (see Table 18.6 for all studies providing evidence from Europe). This might be interpreted as indicating that health responds to changes in permanent income but not to transitory income shocks. As such, it is consistent with the argument made above that sudden income surprises observed over a short period may not provide variation in economic circumstances that is relevant to the determination of health. Sustained differences in income that influence long term behaviour seem more relevant to the evolution of health. However, caution is called for since it is not possible to separate the effect of individual income averaged over a panel from that of time invariant correlated unobservables.

Frijters et al (2005) exploit the largely exogenous income variation generated by the reunification of Germany in 1990 that resulted in sudden large income gains to virtually the whole population of the former East Germany. Reverse causality cannot be eliminated because the East German component of the panel only started in 1990 and so reunification cannot be used as an instrument. Fixed effects models of reported health satisfaction reveal positive effects of income on health in the West but, surprisingly, in the East, where the income variation was much greater, these effects are only observed for males. However, all
estimated effects are very small. Taking into account that the estimates are potentially upwardly biased by the failure to eliminate reverse causality, this study suggests that income does not have a substantial causal impact on health (satisfaction) in Germany.

Using data aggregated at the level of birth cohorts, Deaton and Paxson (2001) find strong negative effects of income on all-cause US mortality in the period 1976-1996. The effects appear strongest in middle age and in young men. But these findings are not uncontroversial. It is difficult to rule out reverse causality in cohort models and the authors’ use of education as an instrument for income is easily criticized. Moreover, the same authors do not find any coherent or stable effects of cohort income on cohort mortality in England and Wales (1971-1998) (Deaton, Paxson 2004). They conclude that the observed correlated cohort income growth and mortality decline in both countries does not necessarily reflect a causal effect of the former on the latter but more plausibly arises from technological advances and the emergence of new diseases, such as AIDS, that affect age groups differentially. In this case, the main identifying assumption of the cohort approach – that age effects on mortality are constant through time – is invalid. This rather negative conclusion has not kept others from adopting a similar approach. Adda et al (2009) study the health effect of permanent income innovations arising from structural changes in the UK economy in the 1980s and 1990s that are assumed to be exogenous. They find that cohort incomes have little effect on a wide range of health outcomes, but do lead to increases in mortality: a 1% increase in income is estimated to lead to 0.7-1 more deaths per 100,000 persons among the prime aged (30-60) population in any given year. This result is in sharp contrast to Deaton and Paxson’s finding of no mortality effect for the UK, and a negative effect of income on mortality for the US. The authors point out that their finding is consistent with Ruhm (2000, 2003), but the latter estimates the effect of transitory income changes arising from the business cycle, while they focus on more permanent income shocks.
Identification of the health effect of windfall gains in income or wealth is rather more transparent and has been a popular strategy adopted in recent studies. The reasoning is that because prizes, lottery wins, investment returns or inheritances are unanticipated, they are more plausibly exogenous to the evolution of health. While this might be true, one may question the relevance of windfall gains to understanding the large differences in morbidity and mortality between the rich and the poor that are likely to arise from sustained differences in health behaviour, and perhaps access to medical care, over many years.

Smith (2007) exploits large wealth gains accumulated by US stockholders during the stock market runups of the late 1980s and 1990s to estimate effects on the onset of major and minor chronic conditions, while conditioning on baseline health, income and wealth. He does not deal with unobserved heterogeneity and so uses the language of prediction, not causation. Wealth changes (positive or negative) do not predict health changes. Using the same PSID data but instrumenting wealth by inheritances, Meer et al (2003) also find no significant effect on health. The same negative result emerges from three studies that test for a response of health to inheritance induced changes in wealth using data on older (50+) individuals from the HRS (Michaud and van Soest 2008, Kim and Ruhm 2012, Carman 2013). Allowing for a rich lag structure and unobserved heterogeneity, Michaud and van Soest (2008), as was noted in section 3.7, find a significant effect of health on wealth, but they find no evidence of a causal effect of (contemporaneous or lagged) wealth on either SAH or chronic conditions. Carman (2013) finds that health is only correlated with inheritances that are anticipated, the exogeneity of which may be doubted.

Perhaps surprisingly, a few European studies do find positive health effects resulting from lottery wins. Using a Swedish panel and instrumenting a measure of permanent income (average income over 15 years) with average lottery winnings, Lindahl (2005) estimates that

39 Inheritances are only used as an instrument for wealth in the models that test for a contemporaneous effect.
an income increase of 10 percent generates a fall in morbidity and a rather spectacular 2-3 percentage point decrease in the probability of dying within 5-10 years. One may be sceptical of the credibility of such a large effect, which exceeds even the raw correlation between income and mortality. Using British data, Gardner and Oswald (2007) find that two years after a win of between £1000 and £120,000, the GHQ index of mental health increased by 1.4 points, on a scale of 36 points. The effect is only significant for males and, surprisingly, for higher income individuals. Using a few more waves of the same data, Apouey and Clark (2013) find that lottery winnings have no significant effect on SAH, but a large positive effect on mental health.

[Table 18.6 here]

4.2.3 Health effects of economic losses

Most of the studies reviewed in the previous subsection look for a health effect of a gain in income or wealth. One might wonder whether the effect is symmetrical. It need not be. For example, winning the lottery may moderately raise mental well-being, but a substantial loss of income may provoke depression or, in some health systems, inability to afford needed medical treatment. The health consequences of economic losses are particularly relevant in the context of the severe recession that has been experienced in many high income economies since 2008.40

McInerney et al (forthcoming) exploit variation in interview dates of the 2008 HRS around the October stock market crash to assess the impact of large wealth losses on mental health (see Table 18.5). The crash reduced wealth and increased depressive symptoms, as well as the

---

40 There is striking and persistent evidence from analyses of aggregated data showing that mortality tends to follow the business cycle, increasing during booms and declining during recessions (Ruhm 2000, Ruhm 2003). While this evidence is not necessarily directly relevant to explanation of the strong cross-sectional positive correlation between income and health, it does appear inconsistent with income gains causally raising health. That recessions are good for health has been contested in research covering the post-2007 Great Recession period by McInerney and Mellor (2012), Stevens et al (2011), Tekin et al (2013), as well as Ruhm (2013) himself.
use of anti-depressants. The effects are non-trivial: for instance, a loss of $50,000 of non-housing wealth is estimated to increase the likelihood of feeling depressed by 1.4 percentage points (8% in relative terms). In contrast to the evidence from the HRS that shows no improvements in health from (inheritance induced) wealth gains, this study suggest that sudden wealth losses can cause immediate and substantial declines in mental health. Whether there is subsequent recovery of mental health status as result of adaptation, or perhaps the onset of physical illness provoked by sustained exposure to depleted economic circumstances, are questions that will hopefully be addressed in the coming years.

In a quite different context in terms of both the economy and the groups affected, Jensen and Richter (2004) study the effect of losses in pension income in Russia during a major crisis period (1995-1996) (Table 18.6). Delayed pension payments had a dramatic impact on living standards, with income declining by up to 24% and poverty rates tripling to over 50 percent. For males, the loss of pension income increased the likelihood of death within two years by 5.8 percent points, and raised functional impairment (ADL) and the probability of experiencing chest pain. These effects are likely to have materialized from substantial and significant reductions in both calorie and protein intake, as well as reduced use of medication for chronic conditions and preventive checkups. There were no effects on women’s health or mortality.

Back in the US, Snyder and Evans (2006) report evidence suggesting that reduced pension income raises health (Table 18.5). They exploit a notch in Social Security payments that resulted in those born after January 1, 1917 receiving sharply lower retirement incomes than contemporaries with identical earnings histories born slightly earlier. There was little time to adjust to the income loss since the legislative changes happened late in working lives. In any case, most of those affected did not realize the impact of the changes until after retirement. The authors find that the decrease in pensions actually reduced mortality and rationalize this
surprising result by a claimed positive health effect of increased postretirement (part time) work effort in response to the income loss.

The contradictory evidence from the US and Russia is most plausibly attributable to differences in the level of income around pensions fluctuated. There is evidence of large mortality reductions (particularly for the poor) resulting from increases in pension paid to US Union Army Veterans at the beginning of the twentieth century (Salm 2011), when both incomes and health were obviously much lower than in the period studied by Snyder and Evans.

4.2.4 Negative health effects of income payments

There is accumulating evidence from the US of health deterioration coinciding with, and even provoked by, the payment of income. The relevant studies do not examine the health response to the level of income, but rather to the time that income is received. They tend to find reductions in health immediately following the receipt of income, suggesting that cash in hand is used to consume health damaging goods.

Evans and Moore (2011) find that mortality increases immediately following the arrival of monthly Social Security payments, regular wage payments for military personnel, tax rebates, and dividend payments. The increase in mortality is large, and occurs for many causes of death connected to short-term behaviour – like heart attacks and traffic accidents – but not for cancer deaths, which suggests that the effects derive from increased risky behaviour. For example, daily mortality of seniors is half a percentage point higher in the week after Social Security pay checks arrive compared to the week before. Mortality in younger populations is even more responsive to income receipt. Dobkin and Puller (2007) find elevated drug-related hospital admissions (23%) and within-hospital mortality (22%) in California in the first few
days of the month for recipients on federal disability insurance programs paid on the first of the month.

Health deterioration in response to payment of a given level of income is not necessarily inconsistent with health improvement arising from a permanently higher level of income. Higher income may afford both a smoother consumption profile and a lifestyle that is freer of health threatening binging on alcohol or drugs. While the evidence rightly makes one wary of the health consequences of increased intermittent cash payments to certain groups, it tells us nothing of how the level of income impacts on health.

4.3 Income and wealth effects on health behaviour

The evidence reviewed in the previous section does not support a strong, or even any, causal effect of income, or wealth, on health. But this may simply reflect the difficulty of observing, even in moderately long panels, the health consequences of changes in health behaviour and utilization of medical care that may only materialize in the long term. In this and the following section, we take it as given that there are health effects of smoking, drinking, etc. and of forgone effective health care, and examine whether there is evidence that economic circumstances impact on these health determinants.

Particularly in the US, more affluent individuals are generally less likely to smoke, drink heavily, be overweight, or use illegal drugs, and are more likely to exercise and engage in preventive care (Cutler, Lleras-Muney 2010, Cutler, Lange et al. 2011). But simple correlations obviously tell us nothing about the presence or direction of causality. Some of the evidence reviewed by Cawley and Ruhm (2011) shows that income and/or wealth increases consumption of tobacco and alcohol. This holds for the response to income shocks in the UK captured by cohort income (Adda, Gaudecker et al. 2009) and lottery winnings (Apoey,
Clark 2013) (see Table 18.6). In the US, Kim and Ruhm (2012) find the wealth gains from inheritances only raise moderate drinking and have no effect on smoking (Table 18.5).

The evidence on income and wealth effects on obesity is mixed but certainly does not support a strong causal effect in either direction. Kim and Ruhm (2012) find some indication of wealth gains reducing the likelihood of being overweight, which is consistent with Swedish evidence based on lottery winnings (Lindahl 2005). Cawley et al. (2010) use US National Health Interview survey data and the Social Security notch as an IV for income and find no impact of income on weight or obesity. Schmeiser (2009), exploiting variation across US states in the generosity of the Earned Income Tax Credit (EITC), finds no effect of income on weight for men, and a positive effect for women: an additional $1,000 per year is associated with a gain of no more than 1.80 pounds (0.82 kilos).

Galama and van Kippersluis (2010) extend Grossman’s (1972) health capital model with the aim of understanding how health behaviour may differ by wealth. They distinguish between healthy consumption, which reduces the rate of depreciation of health e.g. good housing, vitamins, muesli, and unhealthy consumption, which increases health depreciation e.g. cigarettes, excessive alcohol, etc. Wealth has a positive effect on healthy consumption both because of a pure wealth effect and because higher wealth raises health investment that is assumed to exhibit diminishing returns with the result that depreciated health is more expensive to replace through medical care. The effect on unhealthy consumption is ambiguous since the wealth and ‘price’ effects go in opposite directions. The wealthy are less inclined to run down their health because of the higher marginal cost of replacement. The model predicts that, under arguably plausible assumptions, the wealthy will be more likely to

---

41 The assumption that health investment technology exhibits diminishing returns is a departure from Grossman’s model that assumes constant returns. With diminishing returns, since greater wealth raises the demand for health and the level of investment, the marginal cost of producing a unit of health by investing in medical care is greater at higher levels of wealth. This higher marginal cost gives the wealthy an added incentive to look after their health.
engage in moderately unhealthy consumption (wealth effect dominates) and less likely to partake in severely unhealthy consumption (price effect dominates).

Van Kippersluis and Galama (2013) test these predictions with wealth gains instrumented by lottery winnings in British (BHPS) data (Table 18.6), as in Apouey and Clark (2013), and by inheritances in US (HRS) data (Table 18.5), as in Kim and Ruhm (2012). Unlike the earlier studies, they use fixed effects models to deal with unobserved heterogeneity and find robust evidence that wealth increases the probability of drinking alcohol, but it has no effect on the number of drinks and heavy drinking. This is consistent with the direct wealth effect dominating for behaviour that is moderately unhealthy, if taking any alcohol can be considered unhealthy at all. Their results for smoking are less conclusive: a lottery win in the UK does not increase smoking, which is inconsistent with Apouey and Clark (2013), but inheritance receipt in the US does immediately increase both the prevalence and intensity of smoking, which is inconsistent with Kim and Ruhm (2012) – who look at longer term effects – and also with the prediction of the theory.

Income opens consumption opportunities. For unhealthy consumption to explain income-related inequality in health, tobacco, alcohol, fatty foods, etc. would have to be grossly inferior goods. There is no evidence of this. This is not to say that health behaviour is not an important contributor to the social, as opposed to economic, gradient in health. On the contrary, health behaviour can account for a large proportion of the differences in health across education groups (Cutler, Lleras-Muney 2010, Cutler, Lleras-Muney et al. 2011). But it is likely to be the preferences and knowledge of higher education groups, and not their wealth, that leads them to adopt healthier lifestyles.
4.4 Income effects on medical care

As pointed out earlier, the potential for medical care to contribute to health differences by income is constrained in many high income countries by the dominance of public health insurance. The income gradient in utilization of medical care should be stronger in countries, such as the US, that give the market greater scope in the financing of the health system. But public health care seldom completely crowds out private care and even within the European social health insurance and national health service systems specialist care is often distributed in favour of the better-off (Van Doorslaer, Wagstaff et al. 2000, Van Doorslaer, Koolman et al. 2004, Van Doorslaer, Masseria et al. 2006). But while the income elasticity of demand for medical care has been the subject of numerous studies, the literature is surprisingly thin if restricted to that capable of delivering evidence of a causal effect on income on utilization.\(^{42}\)

The Kim and Ruhm (2012) study using the US HRS cited above finds that wealth gains from inheritances raise utilization of many types of medical services and out-of-pocket spending. Using a sample of the old (70+) US population whose drug expenses were not covered by Medicare at the time, Moran and Simon (2006) find a large and statistically significant effect of income instrumented by the Social Security notch on prescription drug utilization, though only for households that have low education and do not have high income (<75th percentile) (Table 18.5). Their estimates of the income elasticity are all above 1. Goda et al (2011) extend the analysis to estimation of the impact of income on utilization of long-term care, which is also not fully covered by Medicare. Their counter-intuitive finding is that a positive permanent income shock lowers nursing home use but increases the utilization of paid home care services. The magnitude of their estimates suggests that moderate reductions in post-retirement income can substantially alter long-term utilization patterns. It is important to bear

\(^{42}\) Virtually every textbook in health economics devotes a chapter to the demand for medical care including estimates of income elasticity (e.g. Sloan and Hsieh, 2012). We do not cover the literature on the effect of income on the demand for health insurance, which would take us some distance from the income-health nexus. It is well-known that uptake of insurance is very much income-related.
in mind that the estimated positive income effects on both drug and long-term care utilization pertain only to the elderly who had been low-wage workers because the impact of the Social Security notch on the pensions of older cohorts who had received higher earnings was relatively weak.

These three studies of the older US population confirm what one would expect. Medical care is a normal good. Where universal public health insurance coverage is absent, individuals that can afford more and better health care will purchase it.

4.5 Income effects on child health

In section 3.4, we concluded that early life health conditions have an economically significant effect on economic well-being in adulthood. Currie (2009) proposes that child health is a potentially important contributor to the intergenerational transmission of education and economic status. The idea is that less educated, poor parents are more likely to give birth to and rear less healthy children. Childhood ill-health interferes with human capital acquisition and directly constrains health capital in adulthood, which further reduces earnings potential. A cycle of poverty is propelled by childhood ill-health. Poverty begets childhood illness, which generates poverty later in life. If true, this would give health a role not only in the creation of inequality, through health shocks that increase income dispersion, but in its perpetuation across generations. Whether parental income does constrain child health is therefore an important question to be addressed not only from a health perspective but also from that of economic inequality.

We focus here on the evidence that parental economic circumstances constrain child health and skip consideration of the mechanisms through which an effect may arise. Almond and Currie (2012) use Cunha and Heckman’s (2007) model of investment in cognitive and noncognitive skills of children to provide a framework for thinking about the evolution of
children’s human capital, including health. Further development of this model to formally incorporate health may provide insight into the impact of parental income on child health.

From a review of the evidence, Currie (2009) concludes that while there is little doubt that children from less privileged backgrounds are less healthy, there is insufficient evidence to conclude that, in a high income country context, this arises from a causal effect. Identification of a causal effect of parental income on child health should, nonetheless, be easier than is identification of the (own) income effect on adult health. The reason is that reverse causality is less of an issue since children generally do not earn income, at least in high income countries. This has been one of the main motivations for researchers to examine the impact of parental income on child health. Of course, reverse causality is not entirely eliminated since illness of a child may interfere with his or her parents’ work activity, and correlated unobservables remain a substantial problem.

4.5.1 Evidence from high income countries

The correlation between family income and children’s general health strengthens as children grow older in the US (Case, Lubotsky et al. 2002) and Canada (Currie, Stabile 2003), suggesting that the disadvantages associated with parental income accumulate as children age. The steepening of the gradient with age can be due to poorer children being hit by more health shocks and/or ability to recover from illness being less for poorer children, given constrained access to medical care. In the US, the strengthening of the gradient is due to a combination of these effects (Case, Lubotsky et al. 2002, Condliffe, Link 2008), whereas in Canada, consistent with its universal health care system, it is only due to poor children becoming sick more frequently (Currie and Stabile, 2003).43

43 (Allin, Stabile 2012) find no evidence that health care utilization is an important factor in generation of the gradient in Canada.
These findings are not generally confirmed for other countries. (Khanam, Nghiem et al. 2009) find that there is a gradient in Australia that strengthens with age when similar covariates to those used by Case et al. (2002) are included. However, the gradient disappears when they include a richer set of controls, in particular maternal health, suggesting that there may be no causal effect. (Reinhold, Jürges 2012) find that the parental income gradient in child health in Germany is as strong as it is in the US but it does not steepen as children grow older, which could be attributed to the constraining effect of universal health care.

The UK evidence is mixed, with Currie et al (2007) Case et al. (2008) arriving at different conclusions from analyses of the same survey. Currie et al. (2007) find a significant family income gradient in child general health that increases between ages 0–3 and 4–8 and decreases afterwards. Case et al. (2008) add three years of data and find that the gradient keeps increasing until age 12. Analyses of a rich data set from one region of England reveal a gradient that does not increase between birth and age 7 and almost disappears with an expanded set of controls, including parental behaviors and health (Burgess, Propper et al. 2004, Propper, Rigg et al. 2007) Using a nationally representative sample, Apouey and Geoffard (2013) find a gradient that persists up to the age of 17 but no evidence that utilization of health care, housing conditions, nutrition, or clothing are important mechanisms generating it.

In North America and Europe, children from poorer households are less healthy. Whether this arises from an effect of parental income or some other characteristic of the family associated with both income and child health cannot be established from the studies cited above. Using reform induced variation in the US Earned Income Tax Credit (EITC), Hoynes et al (2012) estimate that increased maternal income reduces the incidence of low birth weight and increases mean birth weight. For single low-educated mothers, an increase of $1000 in the EITC generates a 6.7-10.8% reduction in the incidence of low birth weight (see Table 18.7 for
this and other studies providing causal evidence in this and next subsection). The effect appears to be mediated through slightly greater use of prenatal care and much more substantial reductions in smoking and drinking during pregnancy.\textsuperscript{44} These estimates suggest sizeable gains in infant health from income increases among low income populations.\textsuperscript{45} A much more modest effect is estimated from data on 14 million US births between 1989 and 2004 that uses a census division-year specific index of skilled-biased technological change to instrument mothers’ earnings (Mocan, Raschke et al. 2013). For low educated (i.e. no more than high school diploma) unmarried mothers who are unlikely to be on Medicaid – public health insurance for low income households – increased earnings raise utilization of prenatal care, as well as birth weight and gestational age. The fact that there are no significant effects for births to high educated mothers and to all mothers who are likely to be covered by Medicaid suggests that low income constrains access to maternity care for those lacking insurance cover. However, the effects are very small. A doubling of earnings would raise birth weight by only 100 grams and gestational age by only two-thirds of a week.

[Table 18.7 here]

\textbf{4.5.2 Evidence from low and middle income countries}

One would expect health in general, and child health in particular, to be more contingent on income in low income settings where nutritional needs to sustain health are often not assured and universal health insurance coverage is absent with most medical care being paid for out-of-pocket. Indeed, the economic gradient in health is particularly steep in low and middle income countries and is evident in critical indicators, such as infant mortality (Commission on

\textsuperscript{44} A $1000 credit received by a low-educated single mother is estimated to increase the propensity to use prenatal care by 0.65 percentage points (from a baseline of 96%), and to reduce the likelihood of smoking by 1.2 percentage points (baseline of 30%) and of consuming alcohol by 1.1 percentage points (baseline of 3.3%). It is not clear why increased income reduces smoking and drinking, although one might suppose that it has to do with reduced financial stress.

\textsuperscript{45} There is also evidence of the child health impact of targeted programs like food stamps (Almond, Hoynes et al. 2011) or food and nutrition vouchers (Hoynes, Page et al. 2011). These are not considered here because of their conditional nature, although (Hoynes, Schanzenbach 2009) claim that recipients of food stamps behave as if the benefits were paid in cash.
the Social Determinants of Health 2008). The evidence that the gradient in child health does derive, at least in part, from a causal impact on economic circumstances on health is much more clear-cut than that from high income countries (see Table 18.7 for studies cited in this subsection).  

Duflo (2000, 2003) examines whether the extension of pensions to black South Africans in the early 1990s had an impact on the nutritional status of children. An effect may have been anticipated since more than a quarter of black children under-five lived with a pension recipient by the end of the period studied. The analysis reveals substantial positive effects of pensions paid to women on the weight and height of girls but no significant effects on the nutritional status of boys and no effect of pensions paid to males. The effects are very large. Payment of a pension to a woman is estimated to raise both weight and height of girls by 1.2 standard deviations over a two year period (Duflo 2003). The income gain was also large with pension benefits being around twice the median per capita income in the rural areas at the time. These results suggest that income can have very large positive effects on child health in low income settings but whether this effect materializes crucially depends on who receives the income. Consistent with this, an unconditional cash grant paid to child caregivers (mostly women) in South Africa has been demonstrated to significantly boost child height (Aguëro,

---

46 A number of studies have exploited macroeconomic shocks to identify the impact of income on health and (infant) mortality in the developing world, e.g. in Mexico (Cutler, Knaul et al. 2002), in Peru (Paxson, Schady 2005), in India (Bhalotra 2010), in Colombia (Miller, Urdinola 2010) and in 59 countries (Baird, Friedman et al. 2011). While these studies are able to identify health effects at the individual level, they are unable to trace the income consequences of the macro shocks at this level. They tend to find substantial negative effects of aggregate income on mortality, but tell us little about the extent to which variation in income across individuals generates inequality in health.

47 The effect on weight (-for-height z-score), which should respond immediately to improved nutrition, is identified by comparing children living in households with elderly relatives eligible for pensions (>59 for females and >64 for males) with others with older relatives that did not quite reach the age of pension qualification. The effect on height (-for-age z-score), which reflects longer term nutritional intake, is identified by comparing the height deficits of younger and older children living in households with an elderly person eligible for a pension relative to those in other households. A smaller deficit among younger children is consistent with a positive impact of income on height since the younger children lived in households benefiting from the pension extension for a larger proportion of their lives.

48 The health gains from pension income in South Africa are not confined to children. Case (2004) finds that the extra income brought by the presence of a pension in a household (equal to 2.5 times the median income in the sample analyzed) improves the health status of all adults in households in which income is pooled.
Carter et al. 2009). On the basis of the observed relationship between adult height and earnings, the projected discounted return to the grant is estimated to be as much as 50%.

Evidence of income effects on child health in poor populations obtained from unconditional cash transfer programmes in Latin America is mixed. Ecuador’s Bono de Desarrollo Humano (BDH) pays $15 per month – equivalent to 6-10% of average household expenditure in the target group – to mothers of children below the age of 17 in the poorest two-fifths of the population. It has been found to have no significant impact on the health (height and haemoglobin concentrations) of children aged 1-3 years (Fernald, Hidrobo 2011), and among older children aged 3-7 years there are only modest effects (on haemoglobin and deworming treatments) for the poorest (Paxson, Schady 2010). Uruguay’s PANES programme targets poorer households, restricting payment of the generous monthly cash transfer – equivalent to 50% of average pre-program income for recipient households and up to 100% of income for households with a recent birth – to households in the bottom decile. It is estimated to reduce the incidence of low birth weight by 1.5 percentage points relative to a baseline of 10 percent (Amarante, Manacorda et al. 2011). This effect appears to materialize through improved maternal nutrition, reduced smoking during pregnancy, a large reduction in the proportion of children born to unmarried parents, and a modest reduction in maternal labor supply. The larger health impact relative to that of the general cash transfer in Ecuador is plausibly explained by the greater magnitude of the payment and its direction to (relatively) poorer households.

---

49 Conditional transfer schemes are less interesting for our purpose since payments are made conditional upon behaviour (e.g. school attendance, medical care receipt, attendance of preventative health services, health and nutrition education) that are intended to have a direct impact on health. While many of these programmes have proven to be highly effective Gertler, (2004), Rivera et al (2004), and Fernald, et al (2008), it is difficult to separate the pure income effect from the incentive effect.

50 In rural areas, vitamin A and iron supplementation did increase and language development improved among children aged 1-3 years (Fernald, Hidrobo 2011).
4.6 Conclusion

In this section we set out to determine the extent to which differences in economic circumstances contribute to health inequality across individuals. Do the poor experience worse health because they are poor? Answering this question is difficult because worse health would be expected to be associated with lower income even without being caused by it. This has driven researchers to search for phenomena that generate variation in income or wealth without being caused by or associated with health. In high income countries, this research enterprise has tended to produce evidence indicative of no impact of income (wealth) on health in adulthood, or effects that are small in comparison to the observed income-health gradient, suggesting that the association does not derive from financial resources impinging on health. To an extent, this is a plausible finding. Variation in health arises from differences in the health stock we are endowed with (genetics), the extent to which we look after this endowment (lifestyle and living conditions), opportunities to repair it when it gets damaged (medical care) and luck. Financial resources cannot influence the first and last determinants. Most high income countries offer universal health care coverage irrespective of ability to pay, which greatly weakens the economic impact on the third determinant. That leaves lifestyle and living conditions. Most research concentrates on the former, and, within this, on what we do that is bad for our health, rather than what we do that is good for it. It would be perverse if greater ability to afford indulgencies in unhealthy behaviour, such as smoking and drinking, explained why the better-off are in better health. In fact, the rich tend to lead less unhealthy lives but that is not because they are economically privileged. More likely, it has to do with their education advantage.

The scope for money to buy health in the developed world is limited. Mental health appears to respond to economic circumstances, with losses producing larger deteriorations in mental well-being than gains generate improvements. But there is little evidence that physical health
problems are provoked by worsened personal finances. However, we suspect that there is much that current research is missing.

Identification of the effect of one stock variable (financial wealth) on another (health capital) is far from easy. There is a risk that identification strategies that focus on very local effects of windfall gains from lottery wins, inheritances, or tax/benefit reforms throw away effects that accumulate over the lifecycle together with the bathwater containing the common unobservables. The determinant that is more permanent – living conditions – tends to get overlooked in economics research. This includes housing and features of the built and social environment that vary with the economic status of neighbourhoods, e.g. pollution, leisure facilities, open spaces, food quality and crime. Money can afford improved housing quality and relocation but it takes a very large economic shock to achieve this. Chronic poverty can entail damp walls, confined spaces, disruptive neighbours, polluted air and a threat of violence that gradually, or perhaps suddenly, take a toll on health.

The health experience of the chronically poor has no influence on the estimated impact of wealth on health in some research, such as that identifying off variation in inheritances. Other evidence, such as that based on lottery wins, does potentially capture exposure among the poor but the sudden and often moderate gains in cash involved may not be sufficient to substantially change living conditions and, even if they are, the observation period is unlikely to be sufficiently long to detect impacts on chronic health problems that may only slowly respond to material circumstances. We are hesitant to conclude that lack of evidence of an impact of wealth on adult physical health in much of the developed world means that there is no effect.

One can be more confident that the worse health of poorer children, which is unfortunately still observed in many high income countries, is not simply a reflection of health constraining
earnings. But it may arise from the fact that poorer parents are also less educated and the likelihood that this impacts on child health. As would be anticipated, the strongest evidence that economic conditions determine inequalities in (child) health comes from the developing world. But even here more income does not necessarily bring better health. Money may be able to buy health when nutritional status is low and many cannot afford medical care, but the money must be given to those that value health highly. There is some evidence that women prioritize child health more than men.

5. Economic inequality as a determinant of health

5.1. Overview

More than twenty years ago, Richard Wilkinson (1990, 1992) introduced the hypothesis that income inequality is harmful to health. He showed that countries with higher income inequality have lower life expectancy, and others soon confirmed a negative association with other measures of population health (Waldmann 1992, Steckel 1995). According to one variant of the hypothesis, this cross-country association reflects a causal effect of income inequality on individual health via psychosocial mechanisms: striving to keep up with the Joneses in societies with higher levels of income inequality raises levels of stress. Income redistribution can potentially raise average health not only because of any greater responsiveness of health to income at lower levels of income, if indeed there is a causal effect of income on health, but also because narrower disparities in income are good for everyone’s health, including that of the rich.

Validity of the hypothesis has been heavily debated (Gravelle 1998, Smith 1999, Wilkinson, Pickett 2006). Its recent extension to the claim that economic inequality is responsible for a host of societal ills – violence, teenage pregnancy, obesity, mistrust and high incarceration rates – beyond poor health (Wilkinson, Pickett 2010) has attracted much attention and been subject to a good deal of criticism (Snowdon 2010, Saunders, Evans 2010). In the present
context, establishing whether there is a health cost of economic inequality is crucial to normative evaluation of the income distribution.

We focus here on mechanisms through which income inequality can potentially impact health and pay close attention to whether the empirical analyses are capable of testing the hypotheses. We restrict attention to the impact of inequality on morbidity and mortality, and neglect studies of homicides, for which there is general agreement on the importance of income inequality (Deaton 2003, Lynch, Smith et al. 2004). We steer clear of the happiness literature that has paid a great deal of attention to income inequality (Alesina, Di Tella et al. 2004, Clark, Frijters et al. 2008).

In the next subsection we demonstrate the stylized fact that population health is negatively associated with income inequality and outline the mechanisms via which it is proposed that income inequality threatens the health of all individuals. We then consider alternative theories that can explain the negative association between population health and income inequality at the aggregate level without inequality being a threat to the health of all individuals in a society. We then turn to the evidence.

5.2. Basic hypothesis: inequality threatens everyone’s health

Population health rises with per capita income but at a decreasing rate (Preston 1975). Among high income countries, where this so-called Preston curve flattens out, population health is negatively correlated with income inequality (Wilkinson 1992, Wilkinson 1996, Wilkinson, Pickett 2010). This is illustrated in Figure 18.5 using the same data on life expectancy and a measure of inequality that are employed in a popular publication that advances the inequality hypothesis (Wilkinson, Pickett 2010). These observations have inspired the claim that at

---

51 In these data, the relationship appears to be driven by the low inequality and high life expectancy of Japan and Sweden, and the high inequality and low life expectancy of the US and Portugal. Among the bulk of countries with life expectancy of 78-80 years, there appears to be no relationship with income inequality. The criteria used by Wilkinson and Pickett (2010) to select the countries included in the figure, and to exclude others, have been
high levels of average income further material gains become less effective in the generation of health and income differences assume greater relevance (Wilkinson, Pickett 2010).

Two causal mechanisms through which income inequality may threaten the health of all individuals – rich and poor – have been proposed. The first stresses the importance of public provision of health determining goods, while the second focuses on social capital. Some stress a third mechanism via psychosocial mechanisms mentioned above (Wilkinson 1992). Since this theory does not propose that income inequality affects health throughout the distribution of income, but rather that the health of less well-off individuals suffers because of their relative deprivation, we cover it in the next section.

![Figure 18.5: Life expectancy and income inequality in high income countries](image)

Notes: Data are from Wilkinson and Pickett (2010) who, in turn, took them from UN Human Development Reports. Life expectancy is at birth averaged over males and females in 2004. Income inequality is measured by the ratio of income received by the richest 20 percent of households to the income of the poorest 20 percent averaged over the years 2003-2006.

Public provision of goods would create a pathway from income inequality to individual health if inequality impacts on the provision of goods, such as curative and preventive health care, education and sanitation, that determine health. Income inequality might lead to more
criticized by others (Saunders, Evans 2010, Snowdon 2010), but defended by the authors in the second edition of their book (Wilkinson, Pickett 2010).

52 The literature tends to presume that there is a positive effect of the income level on health.
heterogeneous preferences, which will reduce the average value (and thus the provision) of publicly provided goods (Alesina, Baqir et al. 1999, Thorbecke, Charumilind 2002, Deaton 2003). But income inequality could also lead to increased public provision since a more skewed income distribution will reduce the income of the median voter relative to the mean and increase the redistributive effect of public provision financed from non-regressive taxation (Meltzer, Richard 1981). Deaton (2003, forthcoming) – following Sen (1999) – argues that focus should be on the health consequences of political, rather than, economic inequalities. He notes that in 19th century Britain, and in the US and India in the 20th century, substantial improvements in public health were realized after the extension of political rights.

Social capital – cohesion and trust among citizens (Putnam, Leonardi et al. 1993) – is argued to be a consequence of economic inequality and a determinant of health via social and psychosocial support, informal insurance mechanisms and information diffusion (Kawachi, Kennedy 1997, Kennedy, Kawachi et al. 1998, d'Hombres, Rocco et al. 2010, Ronconi, Brown et al. 2012). This hypothesis has received a fair amount of attention in the literature, but whether this is justified depends on the assumption that income inequality reduces social cohesion while lower social cohesion raising income inequality seems equally plausible.

5.3. **Alternative hypotheses: health responds to absolute or relative income**

A negative association between population health and income inequality could arise from dependence of health on absolute or relative income without this necessarily implying that inequality threatens everyone’s health (Wagstaff, van Doorslaer 2000). The absolute income hypothesis is that diminishing health returns to income at the individual health level explain the negative association between average population health and income inequality at the aggregate level (Rodgers 1979, Gravelle 1998, Gravelle, Wildman et al. 2002, Wildman, Gravelle et al. 2003). If the health-income relationship is concave then an increase in the spread of the income distribution will bring down mean health since the health loss to those
becoming poorer is larger than the health gain to those becoming richer. Income redistribution could raise average population health but this would occur without the health of any individual, given his or her income, being directly affected by the level of economic inequality in society. The literature reviewed in section 4 provides only limited evidence of a causal impact of income on health in high income countries but there is some evidence that the effect is stronger among the poor, and there does appear to be an income effect in low income countries. If there are health returns to income, they would appear to be diminishing. But even if there is no (diminishing) causal effect of income on health, a negative statistical association between average health and income inequality will be observed when there is a concave statistical relationship between health and income across individuals.

According to the relative hypotheses, what matters for health is how one’s income fares relative to others. When there are diminishing health returns to the difference between individual income and some aggregate, such as the mean, there will be a negative association between average health and income inequality. These hypotheses are distinguished – unfortunately, often not explicitly – from the income inequality hypothesis considered above by proposing that income inequality only matters to the extent that it increases the number of individuals who have an income deficit relative to some reference level. It is only the health of these individuals that is claimed to be damaged by inequality. In contrast, the income inequality hypothesis postulates that income inequality is a common factor impacting on the health of everyone.

Health is presumed to be dependent on relative income because of psychosocial effects. It is not so much possession of more material goods that matters for health but stress, depression, anxiety, shame and distrust brought on by observing one’s standard of living to fall short of that enjoyed by others. These emotional responses are claimed to trigger health-damaging psychoneuroendocrine reactions, such as increased levels of cortisol (Wilkinson 1992).
even hypothesised that this psychosocial-biological effect may be hard-wired into humans through our evolutionary experience (Wilkinson 2001). Hunter-gatherer societies were extremely egalitarian and humans might not yet be well adapted to the social inequalities that arose in settlement societies. Consistent with the psychosocial-biological mechanism, experiments have found that manipulating baboons – that also have stable, although obviously much simpler, hierarchical societies – into lower social positions induces stress (Sapolsky 2005). It has been hypothesized that low relative economic status may impact negatively on health through epigenetic responses (Wilkinson, Pickett 2010). Wilkinson and Pickett cite evidence of the maternal nursing behaviour of rats affecting the offspring’s epigenome at a (glucocorticoid) receptor known to regulate stress responses (Weaver, Cervoni et al. 2004). They propose that, if this also occurs among humans, then increased stress and cortisol levels experienced throughout life could be due to early life epigenetic processes. These authors do not elaborate on why exposure to such processes may be related to low (parental) socioeconomic position, although there is evidence of associations between epigenetic differences and socioeconomic status – proxied by income position, occupation, education and housing tenure among adults (McGuinness, McGlynn et al. 2012) and children (Borghol, Suderman et al. 2012). Understanding of the epidemiological implications of epigenetic processes is still limited (Relton, Davey Smith 2012) and it would certainly be premature to presume that they contribute to health differences across socioeconomic environments.

Why psychosocial responses should be confined to inequality in the income dimension has been questioned (Deaton 2003, Deaton forthcoming). It seems quite conceivable that they might be triggered, perhaps more strongly so, by comparisons of occupation, education, housing, and so on. While describing the mechanism in the previous paragraphs, we (and the

---

53 While still very much in its infancy, epigenetics is the field of medical science that studies (possibly heritable) random or environment-induced changes in gene expression that are not driven by changes in the underlying DNA sequence (Ebrahim 2012).
authors cited) have often resorted to terms such as social hierarchy. The most influential research on health and social position – the Whitehall studies (Marmot, Rose et al. 1978, Marmot, Davey Smith et al. 1991) – has used occupational grade as the discriminating indicator. The evidence obtained from animal studies cannot, of course, inform of effects in the income dimension. In reviewing the evidence below, we will focus on the health effects of relative income, but this is because of the context of this chapter and not because we believe other dimensions of socioeconomic status to be of lesser importance to health.

Less frequently cited than the psychosocial mechanism is the idea that pecuniary externalities, arising from the pricing of health enhancing goods, reduce average health in regions that are more unequal (Miller, Paxson 2006). Take the case of healthy foods. When the quality and availability of healthy food is comparable across regions but it is more expensive in rich areas, then poor individuals in these areas will have worse health than their equivalents in poorer areas. Health depends negatively on the individual’s income deficit from the regional average because of the price effect on the cost of maintaining health. But there could be an offsetting effect through collectively and locally financed health enhancing goods, which may include some medical care. The larger tax base of wealthier neighbourhoods will increase the supply and quality of health care, which will raise the health of a poor person in the rich region compared with his equivalent in a poor region.

There are three variants of the relative hypothesis – relative income, relative deprivation and relative position – distinguished (again, often not explicitly) by the functional form linking health to income differences (Wagstaff, van Doorslaer, 2000). The relative income hypothesis proposes that the magnitude of the difference between income and that of a reference group is what ultimately matters for health (Deaton 2001a, Deaton 2001b, Deaton 2003, Deaton, Paxson 2004). Average income of the group is mostly used as the point of reference, but other aggregations seem equally plausible and there is no theoretical guidance on this. The
reference group is likely to be unobservable or, at best, observed with error, and this will lead to income inequality re-entering the picture even when it exerts no causal effect on individual health (Deaton 2001b, Deaton 2003). The *relative deprivation* hypothesis posits that health is responsive to the difference between income and all larger incomes within the same reference group (Yitzhaki 1979, Deaton 2001b, Eibner, Evans 2005).\(^{54}\) Lower incomes are assumed to be irrelevant for health and so the point of reference is individual-specific. The *relative position* hypothesis suggests that the magnitude of income differences is unimportant and that health responds only to rank in the income distribution. This hypothesis is closest to the aforementioned theories that stress the importance of social hierarchy. It might also be used to justify choice of a rank-based measure of income-related health inequality, such as the concentration index (Wagstaff, Paci et al. 1991).

### 5.4. Evidence

#### 5.4.1. Empirical challenges

Tests have been performed using data at three levels of aggregation: country, region and individual. The majority of studies, and most of the early ones, have relied on country level data, while early US studies used state level data. Individual level data is required to discriminate between the five hypotheses (income inequality, absolute income, relative income, relative deprivation and relative position) since income inequality will correlate with average population health under all them (Wagstaff, van Doorslaer 2000, Mackenbach 2002, Deaton 2003, Lynch, Smith et al. 2004).

The relative hypotheses provoke many unanswered questions with respect to the reference groups: how are they formed, does each individual have a unique reference, and how how are they to be defined in data? Testing is further complicated by the potential for position in relation to the reference to be endogenous through choice of group membership. Data quality

\(^{54}\) Gravelle and Sutton (2009) also study the opposite situation of individuals caring about being richer than others.
and the reliability of measures of income inequality is another major issue (West 1997, Deaton 2003). Estimates of income inequality or relative income at town or village levels may be derived from relatively few observations and so lack precision (Leigh, Jencks et al. 2009).

A major problem for analyses identifying from cross-country or regional variation is that time-invariant unobservable determinants of health may be correlated with income inequality. Fixed effects methods are unlikely to prove successful at aggregated levels since income inequality tends to evolve rather slowly and measurement error bias is compounded (Babones 2008). In addition, fixed effect estimators only identify short run effects and may fail to detect inequality effects operating with a lag. Reverse causality is unlikely to be a major problem at more aggregated levels, but through one or more of the mechanisms identified in section 3, relative income could certainly be a function of health. This would tend to induce bias toward concluding that low relative income exerts a negative impact on health. These limitations must be kept in mind in interpreting the evidence.

We differentiate the evidence by the nature of the hypothesis tested and the level of data aggregation. Given that the impact of income on health is covered in section 4, we do not explicitly consider evidence on the absolute income hypothesis, although we do note what happens to the health-income inequality relationship when individual income is controlled for. Studies published since previous reviews and a few key earlier papers are summarised in Tables 18.8-18.10.

5.4.2. Income inequality hypothesis

Previous reviews have concluded that the evidence does not point to income inequality as an important determinant of individual health, and this seems to hold for both morbidity and mortality (Wagstaff, van Doorslaer 2000, Deaton 2003, Lynch, Smith et al. 2004,
Subramanian, Kawachi 2004, Leigh, Jencks et al. 2009). They also infer from the literature that state-level income inequality associates negatively with health in the US, but this is not true in other countries and this most likely reflects racial composition at the state level, although this interpretation is disputed (Subramanian, Kawachi 2004). There is agreement on the importance of appropriately defining reference groups: when references are defined with respect to smaller geographic units, such as towns or cities, there is less evidence of an association between income inequality and health.

Cross country data

Cross-country positive association between mortality and income inequality is well documented and has most recently been confirmed for 12 European countries (Pascual, Cantarero et al. 2005, Cantarero, Pascual et al. 2005), 25 high income countries (Wilkinson, Pickett 2010), and across many countries worldwide (Babones 2008, Tacke, Waldmann 2013) (Table 18.8). But the the positive association between infant mortality and income inequality across OECD countries evident in 1995 was no longer apparent in 2005 data (Regidor, Martínez et al. 2012). It also appears that the relationship between population health and income inequality is reversed across low income countries (Nilsson, Bergh 2013).

The association between population health, measured by life expectancy or infant mortality, and income inequality that is observed across countries is not evident in time variation within countries, even when exploiting very long time series. For example, there is no association between income inequality and mortality between 1900 and 1998 in the US (Lynch, Smith et al. 2004). Data for Australia, Canada, New Zealand, the US, and 8 European countries display no associations between the within-country evolution of either life expectancy or infant mortality and the income share of the richest ten percent (contemporaneous and lagged)

Wilkinson and Pickett (2006) dissent referring mainly to studies using country level or state level analyses to infer that income inequality is important for health.
between 1903 and 2003 (Leigh, Jencks 2007). There is also no relationship in data for 22 Latin American countries between 1960 and 2007 (Biggs, King et al. 2010). There is an association between the change in income inequality, on the one hand, and the change in life expectancy or infant mortality, on the other, for over 90 countries between 1975 and 1995, but this disappears when the change in GDP per capita is controlled for (Babones 2008). Finally, Avendano (2012) finds no association between within-country variation in infant mortality and income inequality over four decades for 34 OECD countries. This finding remains unchanged after allowing for country specific (linear) time trends or allowing for a lag of 15 years between changes in income inequality and changes in infant mortality.

The absence of any evidence that population health moves with changes in income inequality strongly suggests that the static cross-country relationship does not derive from a causal effect of income inequality on health.

[Table 18.8 here]

**Regional and cohort level data**

Conditional on average regional income, life expectancy is sometimes negatively associated with regional income inequality. The negative relationship holds across US states but not at lower levels of aggregation (metropolitan areas, cities)\(^{56}\) and the evidence is mixed in other countries (Wagstaff, van Doorslaer 2000, Deaton 2003, Lynch, Smith et al. 2004, Subramanian, Kawachi 2004, Wilkinson, Pickett 2006, Leigh, Jencks et al. 2009). The state-level association in the US could either indicate a causal mechanism or greater aggregation bias at that level. Inclusion of state level variables eliminates (or dramatically reduces) the association, although these may be mediators rather than confounders (Subramanian, Kawachi 2004, Wilkinson, Pickett 2006). Racial composition of US states, control for which knocks

\(^{56}\)There is a significant US county level association between income inequality and mortality (Yang, Chen et al. 2012).
out the effect of income inequality, might itself be related to the provision and quality of publicly provided health care (Ash, Robinson 2009, Deaton, Lubotsky 2009) (Table 18.9). That state level income inequality loses much of its explanatory power for mortality after conditioning on measures of social cohesion and interpersonal trust has been interpreted as indicative of a mechanism operating through social capital (Kawachi, Kennedy et al. 1997). Missing from this argument is evidence of a causal effect of income inequality on social capital.


(Table 18.9 here)

**Individual level data**

There is little support for an association between *individual* mortality and income inequality conditional on individual income. This holds both in studies exploiting regional variation in income inequality, that risks confounding from regional health effects, and studies exploiting within country time variation in income inequality, that risks confounding from time trends (Wagstaff, van Doorslaer 2000, Deaton 2003, Lynch, Smith et al. 2004, Subramanian, Kawachi 2004). Morbidity measures, which are more commonly available at the individual level, also have no association with income inequality outside the US. There is evidence that income inequality at the state level, but again not lower levels, is negatively correlated with physical and mental health of the poorest individuals in the US. But this could simply result from state level income inequality picking up the effect of state level differences in public policies towards the poor (Mellor, Milyo 2002) (Table 18.10).

Analyses of individual level (pseudo) panel data find no association between mortality or morbidity and income inequality. Using high quality Swedish administrative data with more
than 10 years of follow up on income and vital status, Gerdtham and Johannesson (2004) find no effect of income inequality measured at the municipality level on mortality after conditioning on individual income and average municipality income. There is also no relationship of income inequality with SAH revealed by analyses of 12 years of British panel data allowing for unobserved heterogeneity (Lorgelly, Lindley 2008) and 22 years of British repeated cross-section data (Gravelle, Sutton 2009). Neither is mental health correlated with income inequality in Australian panel data (Bechtel, Lordan et al. 2012).

Cross-country studies of individual level data largely corroborated the negative finding. There are statistically significant, but economically negligible, effects of regional and nationwide income inequality on SAH in panel data on 11 European countries between 1994 and 2001 (Hildebrand, Kerm 2009). Combining micro data from Australia, Germany, UK, and US with country-level tax records on the income share of the richest percentile (Atkinson, Piketty et al. 2011), Lillard, Burkhauser et al (2012) find that a higher income share of the rich is associated with worse SAH, but once time trends are accounted for the pattern reverses or disappears. The authors also find no evidence that income inequality during the first 20 years of life impacts on current SAH.

Pooling cross-section data for 21 countries, and subject to imposing the same relationship between health and individual income across all those countries, Karlsson, Nilsson et al (2010) find that health is negatively correlated with income inequality in rich income countries, but there is no relationship in middle and low income countries. However, with point-in-time cross-country variation, one can never be sure that the income inequality effect is distinguished from health variation across countries for any other reason. In China, SAH has been found to be positively associated with community-level income inequality at relatively low levels of inequality but the association turns negative at higher inequality (Li, Zhu 2006). Correcting for unobserved heterogeneity using panel data, the negative
relationship persists and is stronger for poorer individuals (Fang, Rizzo 2012). While consistent with the income inequality hypothesis, it is possible that this finding is driven by non-linearity between income and health that is not fully captured in the specification adopted. Child undernutrition in a much lower income country – Zambia – has been found to be negatively correlated with economic inequality (Nilsson, Bergh 2013).

If there is health-related migration across regions with differing levels of inequality, perhaps because of differences in medical care, then estimates from the regression of individual health on regional inequality will be biased. One study avoids this by using the random assignment of refugees to a first area of residence in Sweden (Grönqvist, Johansson et al. 2012). Despite the fact that the range of income inequality across years and municipalities is – perhaps surprisingly – of similar magnitude in Sweden as it is in the US or the UK, hospitalizations, sickness leave and mortality were all found to be unrelated to municipality income inequality. Since refugees are likely to be much poorer and to have quite different references from the general population, one may doubt whether this analysis reveals much about either the income inequality or relative hypotheses that can be generalized.

Finally, some studies find that higher levels of social capital are associated with better individual health (Petrou, Kupek 2008, van Groezen, Jadoenandansing et al. 2011) and others confirm this when using instruments to deal with potential endogeneity of measures of social capital (d'Hombres, Rocco et al. 2010, Ronconi, Brown et al. 2012). None of these studies test whether social capital itself responds to income inequality.

[Table 18.10 here]

5.4.3. Relative hypotheses
The only study that seeks to test all the relative hypotheses (income, deprivation and position) rejects all three in favour of the absolute income hypothesis as an explanation of variation in
SAH using UK longitudinal data (Lorgelly, Lindley 2008). Inconsistent with the relative income hypothesis, conditional on individual income, mortality is found to be lower among individuals living in Swedish municipalities with higher average incomes (Gerdtham, Johannesson 2004). In contrast, mortality risk, especially for black males, is positively correlated with community average income in the US (Miller, Paxson 2006). Data on SAH from 11 European countries are consistent with the relative income hypothesis for males, although the negative correlation of health with average regional income is very small in magnitude, but women, if anything, report slightly better health when regional income is higher (Hildebrand, Kerm 2009). There is little evidence in favour of the relative income hypothesis in UK data on SAH (Gravelle, Sutton 2009).

Analyses of data from Australia (Bechtel, Lordan et al. 2012), China (Li, Zhu 2006) and the UK (Jones, Wildman 2008, Gravelle, Sutton 2009) find little or no evidence consistent with the relative deprivation hypothesis. Although one of the UK studies finds that mental health does fall slightly with relative deprivation (Jones, Wildman 2008), which is consistent with other evidence that mental health, but not physical health or longevity, is negatively associated with relative deprivation (Adjaye-Gbewonyo, Kawachi 2012). One analysis of US data on mortality and SAH does find evidence consistent with the relative deprivation hypothesis when reference groups are defined narrowly (based on race, state, education and age) rather than more broadly (state only) as is the case with most other studies (Eibner, Evans 2005).

5.5. Conclusion

The claim that income inequality is harmful to health has provoked much empirical research. This has delivered little credible evidence to support the hypothesis that income inequality

57 There is a significant negative relationship in some Australian data but the magnitude is negligible (Bechtel, Lordan et al. 2012).
impacts negatively on the health of all individuals in society. Average population health is negatively associated with income inequality across high and middle income countries, but there is no association through time or across regions within countries, except in the US where state-level differences in health seem to be related to racial composition and possibly also social capital. Individual level data on morbidity and mortality from high income countries display no significant, non-negligible relationship to income inequality after controlling for individual income. Few studies have designs capable of testing the hypotheses that relative income, deprivation or economic position causally impact on health. The evidence that exists shows little support for these hypotheses, except maybe for a negative impact of relative deprivation on mental health. But this research has paid insufficient attention to separating relative from absolute income effects, to the strong possibility that health determines relative economic status, and to the appropriate definition, and potential endogeneity, of reference groups.58

The lack of evidence might reflect a lack of well-defined theory and, consequently, precision in the way in which empirical analyses relate to hypotheses. Several potential mechanisms – public provision of goods, social capital, psychosocial mechanisms, pecuniary externalities – have been proposed, but all lack a precise description of how income inequality and/or relative income impact on health. Distinction between the hypotheses is not clear – the relative hypotheses are often claimed to imply an effect of income inequality on the health of all individuals, and it is not clear whether the three relative hypotheses are intended to derive from distinct mechanisms. Psychosocial effects are often loosely cited as the main potential mechanism without specification of how, and for whom, relative economic status provokes stress. Further, it is not clear why priority should be given to relative income, as opposed to some other dimension of socioeconomic position, as a cause of psychosocial stress.

58 Some studies have allowed respondents to define their own reference groups (Theodossiou, Zangelidis 2009, Karlsson, Nilsson et al. 2010, Mangyo, Park 2011), although this introduces an obvious endogeneity.
A cheap call for research designs capable of identifying the impact of income inequality on health would be unhelpful. Sound identification of a causal effect of income on health is difficult enough. Obtaining exogenous variation in income inequality is an even more daunting task. Rather than pursuing further the search for a significant effect of income inequality on health, a more fruitful research agenda would be to directly investigate the causal mechanisms through which it is claimed that health may be related to income inequality. For example, studies have shown that individual health and social capital are associated, but whether this is the result of causality has not received sufficient attention.

6. Conclusion

This chapter has examined three propositions: health differences generate income inequality; income differences generate health inequality; and, income inequality damages health. Grossly simplifying a host of arguments and a vast body of evidence, our verdicts on these three charges are “guilty”, “not proven” and “not guilty” respectively. Appropriately more cautious assessments of the weight of evidence are provided in the conclusions to sections 3, 4 and 5 respectively. Rather than repeat the arguments that lead us to these conclusions, here we restrict attention to their normative and research implications.

Chief among the multitude of mechanisms through which ill-health can impinge on income is the loss of earnings arising from reduced productivity combined with institutional inflexibilities that result in adjustment through employment rather than wages, or marginal changes in work intensity. In high income countries, ill-health is a major cause of labour force withdrawal in middle-age. On pure efficiency grounds, disability insurance is called for to weaken the dependence of income on health and so compress the income distribution. But there is a strong moral hazard effect that makes employment even more sensitive to ill-health.

59 “Not proven” is a verdict available to the courts under Scots law. It is issued when the jury or judge is not convinced of the innocence of the accused but finds the evidence insufficient to prove guilt.
Achieving the optimal balance between income replacement and work incentives is perhaps the greatest challenge for policy that seeks to constrain income inequality arising from ill-health. The task is made even more difficult by increasing economic inequality itself, in the context of which disability insurance can further weaken the labour market attachment of the low skilled facing deteriorating opportunities. Research needs to move beyond identifying the impact ill-health on exit from employment to the design of programmes and incentives that can help individuals experiencing health problems remain in work.

Early life experience is potentially another major route through which health impacts on the distribution of income. Exposure to health risks in utero and ill-health in infancy appear to impact on earnings capacity both by interfering with the accumulation of human capital and skills, and by triggering illnesses in adulthood that disrupt employment. The income distribution observed today is, to some degree, the product of health events that occurred during the childhood of the current adult population. This contribution to economic inequality will be particularly strong if, as appears to be the case, disadvantaged children, who would have grown up to be poorer in any case, face greater health risks. Policies directed at childhood circumstances, including those intended to break the link between parental socioeconomic status and health, may not only be preferred normatively in pursuit of the goal of equality of opportunity (see chapter 5, this volume) but also favoured simply for effectiveness in influencing the distribution of income among adults. However, much of this line of argument is still supposition. The evidence that childhood health is influenced by economic background and determines adult economic outcomes is persuasive but not yet concrete. Fortunately, the pace of progress in this field makes it unnecessary to call for more research on the contribution of early life health to economic inequality.

Our “not proven” verdict on the contribution of income (and wealth) to health inequality arises from the difficulty there would be in detecting an effect if one did, in fact, exist. At
least in high income countries with near universal health insurance coverage and in which the burden of disease is mostly chronic, economic circumstances are likely to exert a toll on health, if at all, over a lifetime. The empirical strategies that have been employed – fixed effects and instrumenting with transitory financial shocks – are incapable of identifying the long run effects that may be operating. Finding random permanent shocks to health from which to estimate the health impact on income is easier than stumbling across exogenous events that permanently change income and allow its effect on health to be identified. The empirical task would undoubtedly be more manageable if there was more theory available to identify precise mechanisms through which income (wealth) might plausibly impact on health. The lack of theory is understandable. Economists are trained to explain the distribution of income, not health. Forty years after Grossman (1972a) introduced the concept of the health production function it remains a black box; all too often cited to motivate study of the relationship of health to some socioeconomic factor but seldom more fully specified to make the mechanism of any effect explicit.

Rather than pursuing further identification of a reduced form effect of income or wealth on health, we believe it is more fruitful to focus on plausible inputs to the health production function that can be influenced by economic status. For example, establishing the health effect of damp, squalid housing is more feasible than finding the health effect of income that affords superior quality housing. This is not merely call for empirical pragmatism. Provided redistribution policy is motivated, in part, by (health) specific egalitarianism – and we attribute the extensive involvement of governments in the provision of health insurance and medical care as being motivated not only by the correction of market failures but also by concern for the distribution of health – then it might be more efficient to enable poor people live in less unhealthy conditions, rather than redistributing cash to them. Once basic
nutritional needs are satisfied and access to medical care has been divorced from ability to pay, the path leading from income to health seems a very long one.

If one switches attention from the distribution of health to that of well-being, then the association between income and health may be used to justify greater redistribution of income, even in the absence of any causal effect. Assuming well-being is increasing in both income and health, the positive correlation between them increases inequality in well-being by more than is implied by the inequality in their marginal distributions (Deaton forthcoming). Redistribution of income toward those in worse health would reduce inequality in well-being both by compensating for sickness and, on average, by reaching poorer individuals (Deaton 2002). According to this argument, redistribution is partially motivated by one dimension of well-being (income) compensating for deficiency in another (health). This is not how health-related income transfers are typically justified. The disabled are paid transfers because their earnings capacity is impaired and/or they have higher costs of living. The transfers are made because ill-health has a causal impact on economic living standards. The ethical argument makes a case for income redistribution to the sick simply because they are sick. Courts awarding damages for injuries irrespective of their consequences for earnings or living costs are consistent with these ethics. But government social policies typically are not. Transfers compensate for financial losses, not reductions in other dimensions of welfare.

With respect to the charge that income inequality threatens health, a case could be made for revising the verdict from “not guilty” to “not proven”. It is fundamentally difficult to separate any potential effect of income inequality on the individual’s health from that of physical environmental, social, cultural or economic determinants of health that operate at the same level as income inequality is measured. Identification of the impact of relative income on

---

60 A still more effective redistribution policy might be one that operates through a factor, perhaps education, that exerts a causal impact on both income and health (Deaton 2002).
health is even more challenging than that of absolute income given the added complexity of defining and measuring the reference point. But the limitations are not only empirical. There is a lack of precision in the theoretical arguments as to why economic inequality should impact negatively on health.

The conclusions offered above are based on the evidence from high income countries. In low income countries, in which a substantial fraction of the population may live close to subsistence and only the economically privileged can afford effective medical care, ill-health is not only an important cause of economic inequality but a consequence of it. But it is the absolute living conditions of the poor, and not their relative deprivation, that takes the toll on health.
### TABLES

Table 18.1. Least squares regressions of log household equivalent income and decompositions of income inequality (relative Gini), United States, the Netherlands and China

<table>
<thead>
<tr>
<th></th>
<th>United States</th>
<th></th>
<th></th>
<th></th>
<th>Netherlands</th>
<th></th>
<th></th>
<th></th>
<th>China</th>
<th></th>
<th></th>
<th></th>
<th>China</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Inequality contribution</td>
<td>OLS coeff.</td>
<td>Inequality contribution</td>
<td>OLS coeff.</td>
<td>Inequality contribution</td>
<td>OLS coeff.</td>
<td>Inequality contribution</td>
<td>OLS coeff.</td>
<td>Inequality contribution</td>
<td>OLS coeff.</td>
<td>Inequality contribution</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Health (SAH) (reference = poor)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>moderate/fair good very good excellent</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>moderate/fair</td>
<td>0.208***</td>
<td>0.146**</td>
<td>0.197***</td>
<td>0.013</td>
<td>0.184***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fair</td>
<td>0.508***</td>
<td>0.235***</td>
<td>0.244***</td>
<td>0.206***</td>
<td>0.152**</td>
<td>0.221***</td>
<td>2.5%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>very good</td>
<td>0.688***</td>
<td>0.314***</td>
<td>0.389***</td>
<td>0.334***</td>
<td>0.232***</td>
<td>0.362***</td>
<td>2.5%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>excellent</td>
<td>0.663***</td>
<td>0.359***</td>
<td>0.359***</td>
<td>0.311**</td>
<td>0.268**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Education (ref. = low)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>middle high</td>
<td>0.445***</td>
<td>0.069***</td>
<td>0.370***</td>
<td>0.390***</td>
<td>0.040**</td>
<td>0.314***</td>
<td>11.1%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>high</td>
<td>0.882***</td>
<td>0.344***</td>
<td>0.754***</td>
<td>0.764***</td>
<td>0.305***</td>
<td>0.641***</td>
<td>11.1%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Gender (ref = female)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>female</td>
<td>0.138***</td>
<td>0.060***</td>
<td>0.021</td>
<td>0.103***</td>
<td>0.036**</td>
<td>0.042**</td>
<td>0.4%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Ethnicity (ref=minority)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>minority</td>
<td>0.440***</td>
<td>0.062**</td>
<td>0.078**</td>
<td>0.392***</td>
<td>0.0043</td>
<td>0.051**</td>
<td>0.3%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Age (ref. = 20-29 years)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30-49 years</td>
<td>0.195***</td>
<td>-0.009</td>
<td>0.048</td>
<td>0.158***</td>
<td>-0.060*</td>
<td>-0.020</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40-49 years</td>
<td>0.392***</td>
<td>0.023</td>
<td>0.097**</td>
<td>0.354***</td>
<td>-0.036</td>
<td>0.023</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50-59 years</td>
<td>0.496***</td>
<td>0.161***</td>
<td>0.106**</td>
<td>0.513***</td>
<td>0.127***</td>
<td>0.017**</td>
<td>3.0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60-69 years</td>
<td>0.585***</td>
<td>0.081**</td>
<td>-0.059</td>
<td>0.617***</td>
<td>0.132***</td>
<td>-0.247**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>70+ years</td>
<td>0.474***</td>
<td>-0.032</td>
<td>-0.098*</td>
<td>0.508**</td>
<td>0.041</td>
<td>-0.311**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Region</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.6%</td>
<td></td>
<td></td>
<td></td>
<td>3.6%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.0%</td>
<td></td>
<td></td>
<td></td>
<td>4.0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Employment status (ref. = employed)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>unemployed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>disabled</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>retired</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>not working</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unexplained (OLS residual)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relative Gini</td>
<td>0.456</td>
<td>0.292</td>
<td>0.472</td>
<td>0.456</td>
<td>0.292</td>
<td>0.472</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of observations</td>
<td>5050</td>
<td>4137</td>
<td>7694</td>
<td>5050</td>
<td>4137</td>
<td>7694</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Notes:** Column headed ‘OLS coeff.’ gives coefficients from least squares regression of log household equivalent income. ‘Inequality contribution’ is the estimated contribution of the factor to inequality in household equivalent income computed from the Shapley value decomposition of the relative Gini index. SAH = self-assessed health. Definitions and means of the dependent variables and covariates are provided in Appendix Table A1. Reference category for ethnicity is not belonging to the main ethnic group, which (white (US), Dutch (NL) and Han (China). Coefficients of regions (US) and provinces (China) are not shown to save space. Region identifiers are not made available with the Dutch data. *,**,*** indicate significant at 10, 5, 1% level.
Table 18.2. Interval regressions of self-assessed health (SAH) and decompositions of inequality (absolute Gini) in predicted SAH, United States, the Netherlands and China

<table>
<thead>
<tr>
<th></th>
<th>United States</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Regression</td>
<td>Inequality</td>
<td>Regression</td>
<td>Inequality</td>
<td>Regression</td>
<td>Inequality</td>
</tr>
<tr>
<td></td>
<td>coeff.</td>
<td>contribution</td>
<td>coeff.</td>
<td>contribution</td>
<td>coeff.</td>
<td>contribution</td>
</tr>
<tr>
<td></td>
<td>baseline</td>
<td>with employment</td>
<td>baseline</td>
<td>with employment</td>
<td>baseline</td>
<td>with employment</td>
</tr>
<tr>
<td>Income quartile</td>
<td>baseline</td>
<td>with employment &amp;</td>
<td>baseline</td>
<td>with employment &amp;</td>
<td>baseline</td>
<td>with employment &amp;</td>
</tr>
<tr>
<td>(reference = poorest</td>
<td>employment</td>
<td>health behaviour</td>
<td>health behaviour</td>
<td>health behaviour</td>
<td>health behaviour</td>
<td>health behaviour</td>
</tr>
<tr>
<td>25%)</td>
<td>25th</td>
<td>44.0%</td>
<td>21.7%</td>
<td>14.3%</td>
<td>34.5%</td>
<td>17.7%</td>
</tr>
<tr>
<td>2nd poorest</td>
<td>0.038***</td>
<td>0.016***</td>
<td>0.032***</td>
<td>0.010**</td>
<td>0.013***</td>
<td>0.016***</td>
</tr>
<tr>
<td>2nd richest</td>
<td>0.075***</td>
<td>0.024***</td>
<td>0.036***</td>
<td>0.022***</td>
<td>0.019***</td>
<td>0.030***</td>
</tr>
<tr>
<td>Education (ref. = low)</td>
<td>0.020***</td>
<td>0.009**</td>
<td>0.015***</td>
<td>0.006</td>
<td>0.011***</td>
<td>0.051***</td>
</tr>
<tr>
<td>middle</td>
<td>0.042***</td>
<td>0.015**</td>
<td>0.022**</td>
<td>0.011**</td>
<td>0.015***</td>
<td>0.054***</td>
</tr>
<tr>
<td>high</td>
<td>0.001</td>
<td>0.003</td>
<td>0.006</td>
<td>0.007</td>
<td>0.004</td>
<td>0.008</td>
</tr>
<tr>
<td>Gender (ref.=female)</td>
<td>0.012**</td>
<td>0.013**</td>
<td>0.013**</td>
<td>0.016**</td>
<td>0.016**</td>
<td>0.016**</td>
</tr>
<tr>
<td>Ethnicity (ref.=minority)</td>
<td>0.025**</td>
<td>0.025**</td>
<td>0.025**</td>
<td>0.025**</td>
<td>0.025**</td>
<td>0.025**</td>
</tr>
<tr>
<td>Age (ref.= 20-29 years)</td>
<td>0.012**</td>
<td>0.021**</td>
<td>0.021**</td>
<td>0.021**</td>
<td>0.021**</td>
<td>0.021**</td>
</tr>
<tr>
<td>30-39 years</td>
<td>-0.012**</td>
<td>-0.012**</td>
<td>-0.012**</td>
<td>-0.012**</td>
<td>-0.012**</td>
<td>-0.012**</td>
</tr>
<tr>
<td>40-49</td>
<td>-0.035***</td>
<td>-0.035***</td>
<td>-0.035***</td>
<td>-0.035***</td>
<td>-0.035***</td>
<td>-0.035***</td>
</tr>
<tr>
<td>50-59</td>
<td>-0.051***</td>
<td>-0.051***</td>
<td>-0.051***</td>
<td>-0.051***</td>
<td>-0.051***</td>
<td>-0.051***</td>
</tr>
<tr>
<td>60-69</td>
<td>-0.054***</td>
<td>-0.054***</td>
<td>-0.054***</td>
<td>-0.054***</td>
<td>-0.054***</td>
<td>-0.054***</td>
</tr>
<tr>
<td>70+</td>
<td>-0.042***</td>
<td>-0.042***</td>
<td>-0.042***</td>
<td>-0.042***</td>
<td>-0.042***</td>
<td>-0.042***</td>
</tr>
<tr>
<td>Region</td>
<td>0.024</td>
<td>0.018</td>
<td>0.021</td>
<td>0.019</td>
<td>0.020</td>
<td>0.021</td>
</tr>
<tr>
<td>Employment status</td>
<td>5050</td>
<td>4137</td>
<td>6794</td>
<td>0.028</td>
<td>0.033</td>
<td>0.038</td>
</tr>
<tr>
<td>Health behaviour</td>
<td>0.031</td>
<td>0.031</td>
<td>0.031</td>
<td>0.031</td>
<td>0.031</td>
<td>0.031</td>
</tr>
<tr>
<td>Absolute Gini</td>
<td>0.024</td>
<td>0.028</td>
<td>0.033</td>
<td>0.014</td>
<td>0.018</td>
<td>0.020</td>
</tr>
<tr>
<td>Number of observations</td>
<td>5050</td>
<td>4137</td>
<td>6794</td>
<td>0.028</td>
<td>0.033</td>
<td>0.038</td>
</tr>
</tbody>
</table>

Notes: Column headed ‘Regression coeff.’ gives the coefficient from the interval regression of SAH with thresholds of categories as defined in footnote 4. ‘Inequality contribution’ is the estimated percentage contribution of the factor to inequality in predicted health (obtained from the respective interval regression) computed from the Shapley value decomposition of the absolute Gini index. The latter takes values between 0 and 0.25, with those bounds indicating minimum and maximum health inequality respectively. Definitions and means of the dependent variables and covariates are provided in appendix Table A1. Reference category for ethnicity as in notes to Table 18.1. Reference category for region/province coefficients not shown. Regression coefficients are shown only for the baseline specification. Extended specifications sequentially add employment status (as Table 18.1) and proxies for health behaviours – smoking and weight. Smoking is measured by an indicator of ever having smoked in US and NL, and currently being a smoker in China. Weight is summarized by a dummy for normal or overweight (18.5<BMI<27.5 for China and 18.5<BMI<30 for US and the Netherlands) and another for obesity (BMI>30 for the US and the Netherlands, and BMI>27.5 for China). Underweight, BMI<18.5 is the reference.
Table 18.3. US evidence of health effects on labour market outcomes

<table>
<thead>
<tr>
<th>Authors</th>
<th>Data</th>
<th>Sample</th>
<th>Health measure</th>
<th>% in ill-health</th>
<th>Biases addressed</th>
<th>Estimator</th>
<th>Wage</th>
<th>Employment</th>
<th>Hours</th>
<th>Earnings/Income</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bound et al (2010)</td>
<td>HRS 1992-98 (4 waves)</td>
<td>Single men 50-62 yrs (baseline)</td>
<td>Latent: SAH instrumented by ADLs</td>
<td>SEL, UH, ME (JB)</td>
<td>DP model by SML</td>
<td>1 std. dev. ↓ from average health @ 60 yrs → Pr(labour force exit) ↑ 8 ppt</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>French (2005)</td>
<td>PSID 1968-1997</td>
<td>20-70 yrs</td>
<td>Work limiting physical impairment or nervous condition</td>
<td>SEL, UH, SIM</td>
<td>FE with SEL &amp; MSM</td>
<td>&lt;35 yrs: no effect, &gt;62 yrs: 3% ↓ 35-62 yrs: 8-17% ↓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Charles (2003)</td>
<td>PSID 1968-1993</td>
<td>Men 22-64 yrs</td>
<td>Work limiting physical impairment</td>
<td>SEL, UH</td>
<td>FE with Heckman selection</td>
<td>42%</td>
<td>46.7%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blau &amp; Gilleskie (2001)</td>
<td>HRS 1992-94 (2 waves)</td>
<td>Men 50-62 yrs</td>
<td>SAH, work limiting disability, major/minor diagnoses, ADLs</td>
<td>SEL, ATT, UH, SIM</td>
<td>FIML</td>
<td>Excellent → poor SAH: Pr(labour force exit) 5.7 ppt, Work limiting disability: Pr(labour force exit) 5.5 ppt</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: Table excludes earlier (pre-1998) studies summarized in tables presented in Currie and Madrian (1999). a Effects on outcomes are with respect to this health measure. b Impact on hourly wage for those in employment. c Impact on probability of employment unless otherwise stated. d Relative impact on annual work hours. e Assuming average annual work hours of 1800. JB=Justification bias, ME=Measurement error, SIM=Simultaneity bias, SEL=Selection bias, UH=Unobservable heterogeneity (OVB bias). See Appendix Table A2 for explanation of other acronyms, including those of datasets and variables.
<table>
<thead>
<tr>
<th>Authors</th>
<th>Country</th>
<th>Data details</th>
<th>Sample</th>
<th>Study details</th>
<th>% in ill-health</th>
<th>Biases addressed</th>
<th>Estimator</th>
<th>Effect of ill-health on:</th>
</tr>
</thead>
<tbody>
<tr>
<td>García Gómez et al (2013)</td>
<td>Netherlands</td>
<td>Administrative 1998-2005</td>
<td>18-64 yrs.</td>
<td>Urgent, unscheduled hospital admission ≥3 nights</td>
<td>0.85%</td>
<td>UH, ME (JB)</td>
<td>DID &amp; matching</td>
<td>↓7.1 ppt</td>
</tr>
<tr>
<td>Halla &amp; Zweimüller (2013)</td>
<td>Austria</td>
<td>Administrative 2000-2007</td>
<td>Private employees 25-50 yrs. 16-64 yrs</td>
<td>Commuting accident ≥1 sick day</td>
<td>0.67%</td>
<td>UH, ME (JB)</td>
<td>DID &amp; matching</td>
<td>↓3.3 ppt</td>
</tr>
<tr>
<td>García Gómez (2011)</td>
<td>9 EU countries</td>
<td>ECHP 1994-2001 (8 waves)</td>
<td>18-64 yrs.</td>
<td>SAH = top 2 → bottom 3. Onset chronic illness/disability ≥3 nights</td>
<td></td>
<td></td>
<td>SEL, UH, ME, SIM</td>
<td>4.8% (daily wage)</td>
</tr>
<tr>
<td>Jones et al (2010)</td>
<td>UK</td>
<td>BHPS 1991-2002 (12 waves)</td>
<td>50-60/65 yrs.</td>
<td>Reported limited activity &amp; SAH instrumented by health problems</td>
<td>SEL, UH (ME (JB) when use instrumented SAH)</td>
<td></td>
<td>2 stage: GOP of SAH → ML of retirement hazard</td>
<td>Effects on retirement hazard: Limited activity ↓5.5 (M), ↓5.8 (F), SAH (excellent → poor/very poor): ↓4.9 (M), ↓7.2 (F) ↓ (large relative to non-health effects)</td>
</tr>
<tr>
<td>Lindeboom &amp; Kerkhofs (2009)</td>
<td>Netherlands</td>
<td>CERRA 1993-95 (2 waves)</td>
<td>Male 43-63 yrs, employed, hhold head 25-64 yrs</td>
<td>Work limitation instrumented by health problems (HSCL)</td>
<td>14%</td>
<td>UH (RE), ME (JB), SIM</td>
<td></td>
<td>↓ (large relative to non-health effects)</td>
</tr>
<tr>
<td>Cai (2009)</td>
<td>Australia</td>
<td>HILDA 2003</td>
<td>Males 25-64 yrs.</td>
<td>SAH (top 3 → bottom 2)</td>
<td>10.3%</td>
<td>SEL, SIM</td>
<td>FIML</td>
<td>↓17-20%</td>
</tr>
<tr>
<td>Disney et al (2006)</td>
<td>UK</td>
<td>BHPS 1991-98 (8 waves)</td>
<td>50-60/64 yrs.</td>
<td>SAH instrumented by health problems &amp; reported limited activity</td>
<td>SEL, UH, ME (JB), SIM</td>
<td></td>
<td>2 stage: Ordered Probit of SAH → FE logit (or RE hazard) of employment DID &amp; matching</td>
<td>↓5 pp</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Personal Income ↓1033</td>
<td></td>
<td></td>
<td></td>
<td>Household income ↓€1927</td>
</tr>
<tr>
<td>Authors</td>
<td>Country</td>
<td>Data</td>
<td>Sample</td>
<td>Health measure</td>
<td>% in ill-health</td>
<td>Biases addressed</td>
<td>Estimator</td>
<td>Effect of ill-health on:</td>
</tr>
<tr>
<td>-------------------------</td>
<td>------------</td>
<td>---------------------------</td>
<td>--------------</td>
<td>-------------------------------------</td>
<td>----------------</td>
<td>------------------</td>
<td>-----------------------------</td>
<td>---------------------------------------------------</td>
</tr>
<tr>
<td>Contoyannis &amp; Rice (2001)</td>
<td>UK</td>
<td>BHPS 1991-96 (6 waves)</td>
<td>16+ finished schooling</td>
<td>SAH, GHQ (psychological ill-health)</td>
<td>SAH &lt; good: Males: 16% Females: 19%</td>
<td>UH</td>
<td>2 stage: FE of work limitation → ML of employment hazard</td>
<td>↓ (ill-health dominant effect on exit thru’ DI)</td>
</tr>
<tr>
<td>Riphan (1999)</td>
<td>Germany</td>
<td>GSOEP 1984-94 (10 waves)</td>
<td>SAH = 45 points on 10 point scale</td>
<td>3.1%</td>
<td>Logit (employment), Unconditional DID (earnings/income)</td>
<td>46 ppt Earnings: growth 41.9 ppt Household income: growth ↓5.2 ppt</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: Table excludes studies published pre-1999. JB=Justification bias, ME=Measurement error, SIM=Simultaneity bias, SEL=Selection bias, UH=Unobservable heterogeneity (OVB bias). See Appendix Table A2 for explanation other acronyms, including those of datasets and variables.

a Effects on outcomes are with respect to this health measure. b Impact on hourly wage for those in employment. c Impact on probability of employment unless otherwise stated. d Estimates from one of estimators used to correct for selectivity bias. Estimates from other estimators vary greatly, reaching effects of 16-19 percentage points.

d Despite the large impact of a health shock on the retirement hazard, few individuals experience such a shock and early retirement is simulated to be only 11% above what it would be in the absence of the shocks incurred.
<table>
<thead>
<tr>
<th>Authors</th>
<th>Data</th>
<th>Sample</th>
<th>Income / wealth measure*</th>
<th>Measures of health / health behaviour</th>
<th>Estimator</th>
<th>IV for income/wealth</th>
<th>Effect on health</th>
<th>Effect on health behavior or medical care use</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>McInerney et al (2013)</td>
<td>HRS 2006 &amp; 2008</td>
<td>Adults 50+</td>
<td>Drop in non-housing wealth</td>
<td>CES-D depression, anxiety, SAH</td>
<td>IV in 1st differences</td>
<td>↑ prob. anti-depressant use for large wealth losses</td>
<td>↑ moderate, not excessive, drinking. ↓ smoking</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Van Kippersluis &amp; Galama (2013)</td>
<td>HRS 1992-2010</td>
<td>50+</td>
<td>Household wealth</td>
<td>Smoking and drinking</td>
<td>IV FE</td>
<td>Interview date post 2008 stock market crash Inheritances</td>
<td>↑ CES-D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kim &amp; Ruhm (2012)</td>
<td>HRS 1992-2006</td>
<td>50-60 at baseline</td>
<td>Inheritances (above and below $10,000)</td>
<td>Mortality, SAH, ADL, CES-D</td>
<td>Discrete time hazard (logit), LPM</td>
<td>↑ 1.4 ppt prob. depressed ↓ prob. of good health</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Salm (2011)</td>
<td>US Vital statistics 1900-1917</td>
<td>Union Army Veterans 70+ at baseline</td>
<td>Social Security (SS) income</td>
<td>Use of home care or nursing home</td>
<td>IV probit SS Notch</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goda et al (2011)</td>
<td>AHEAD 1993-95</td>
<td>55+ Household income</td>
<td>BMI, over- and underweight, obesity</td>
<td>2SLS SS Notch (5-7% income loss)</td>
<td></td>
<td>no effects</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Michaud &amp; van Soest (2008)</td>
<td>HRS 1992–2002 (6 waves)</td>
<td>25-43 years Adults</td>
<td>Household income</td>
<td>Onset of new health conditions / disease</td>
<td>Probit</td>
<td></td>
<td>No effects</td>
<td>Models w/o unobserved heterogeneity or with too few lags → different estimates</td>
<td></td>
</tr>
<tr>
<td>Schmeiser (2008)</td>
<td>NLSY 1979</td>
<td>Low income 25–43 years Adults</td>
<td>Household income, total &amp; stock market wealth</td>
<td>Use of prescription drugs</td>
<td>OLS and IV SS Notch</td>
<td></td>
<td>↓ prescription drug use among low income</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smith (2007)</td>
<td>PSID 1984-1999</td>
<td>Adults</td>
<td>Households income, total &amp; stock market wealth</td>
<td>5 year mortality rates</td>
<td>DID &amp; RDD SS Notch</td>
<td></td>
<td>↑ mortality</td>
<td>No effect of either positive or negative wealth change</td>
<td></td>
</tr>
<tr>
<td>Authors</td>
<td>Data</td>
<td>Sample</td>
<td>Income / wealth measure</td>
<td>Measures of health / health behaviour</td>
<td>Estimator</td>
<td>IV for income/wealth</td>
<td>Effect on health behavior or medical care use</td>
<td>Remarks</td>
<td></td>
</tr>
<tr>
<td>--------------------</td>
<td>-----------------</td>
<td>---------</td>
<td>-------------------------</td>
<td>-------------------------------------</td>
<td>--------------------</td>
<td>---------------------</td>
<td>-----------------------------------------------</td>
<td>---------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Adams et al (2003)</td>
<td>AHEAD 1994-1998</td>
<td>70+</td>
<td>Liquid and non-liquid wealth, household income</td>
<td>Mortality, acute and chronic conditions, SAH, ADL, BMI, smoking</td>
<td>Probit &amp; Ordered Probit</td>
<td>Non-causality not rejected except for mental health and some chronic conditions.</td>
<td>Non causality rejected only for smoking (males) and BMI (females).</td>
<td>UH not eliminated</td>
<td></td>
</tr>
<tr>
<td>Meer et al (2003)</td>
<td>PSID 1984–99</td>
<td>Household heads</td>
<td>Net wealth (not pension wealth)</td>
<td>SAH</td>
<td>Two-stage probit</td>
<td>Gifts or inheritances &gt; $10,000</td>
<td>No effect</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Ettner (1996)</td>
<td>NSFH 1987, SIPP 1986-7, NHIS 1988</td>
<td>18-65</td>
<td>Family income</td>
<td>SAH, work and functional limitations, bed days, alcohol consumption, depressive symptoms</td>
<td>OLS and 2SLS</td>
<td>Work experience &amp; state unemp. rate, parental &amp; spousal education</td>
<td>Strong and large ↑ physical &amp; mental health</td>
<td>No effects Dubious IVs</td>
<td></td>
</tr>
</tbody>
</table>

Notes: * Effect on health / health behaviour is with respect to health measure defined in this column. See Appendix A for explanation of dataset, variable and estimator acronyms.
### Table 18.6: European evidence of income and wealth effects on adult health & health behaviour

<table>
<thead>
<tr>
<th>Authors</th>
<th>Country &amp; data</th>
<th>Sample</th>
<th>Income / wealth measure</th>
<th>Measures of health / health behaviour</th>
<th>Estimator</th>
<th>Effect on health</th>
<th>Effect on health behavior or medical care use</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apouey &amp; Clark (2013)</td>
<td>UK BHPS 1997-2005</td>
<td>Adults</td>
<td>lottery winnings</td>
<td>SAH, mental health (GHQ), physical health problems, smoking, social drinking Smoking and drinking</td>
<td>FE OLS</td>
<td>Mental health ↑ SAH &amp; physical health – no effects</td>
<td>↑ smoking &amp; drinking</td>
<td></td>
</tr>
<tr>
<td>Van Kippersluis &amp; Galama (2013)</td>
<td>UK BHPS 1997 -2008</td>
<td>Adults</td>
<td>Household wealth instrumented by lottery win</td>
<td>Mortality, SAH, chronic illness, blood pressure, cardiovascular /respiratory diseases. Smoking &amp; drinking</td>
<td>IV FE</td>
<td>↑ in moderate, not excessive drinking. No effect on smoking</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gardner &amp; Oswald (2007)</td>
<td>UK BHPS 1996-2003</td>
<td>Adults</td>
<td>Lottery winnings 1000-120000 GBP</td>
<td>Health satisfaction (1-10)</td>
<td>FE ordered logit</td>
<td>↑ mental health (41.4/36 GHQ points) 2 years after win. Significant but small ↑ health; in East Germany is small effect for males only</td>
<td>-</td>
<td>Estimates derived from only 137 wins &gt;1000 GBP. No exogenous income variation for West Germans</td>
</tr>
<tr>
<td>Frijters et al (2005)</td>
<td>Germany SOEP 1984- '02 (West) &amp; 1990-’02 (East) Sweden SLLS 1968, 1974 &amp; 1981</td>
<td>18+</td>
<td>Household income</td>
<td>Health satisfaction (1-10)</td>
<td>FE ordered logit</td>
<td>↑ mental health (41.4/36 GHQ points) 2 years after win. Significant but small ↑ health; in East Germany is small effect for males only</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Lindahl (2005)</td>
<td>Sweden SLLS 1968, 1974 &amp; 1981</td>
<td>Adults</td>
<td>Disposable family income (tax register) instrumented by lottery win</td>
<td>Health index, mortality, overweight</td>
<td>2SLS and prob</td>
<td>10 % ↑ in income ➔ ↑ health index 0.04-0.09 std. dev. &amp; + 2-3 ppt prob. death within 10 years</td>
<td>No effect on overweight IV estimates for mortality vy large. No effects for older (60+)</td>
<td></td>
</tr>
<tr>
<td>Contoyannis et al (2004)</td>
<td>UK BHPS 1991-99</td>
<td>16+</td>
<td>Household income (current=annual permanent= mean over panel) Gross income per adult equivalent (cohort average)</td>
<td>SAH</td>
<td>Dynamic RE Ordered Probit</td>
<td>↑ health larger for permanent than current income than current income. Larger effect for men than women</td>
<td>Cannot separate UH from permanent income in mean income effect</td>
<td></td>
</tr>
</tbody>
</table>

Notes: * Effect on health / health behaviour is with respect to health measure defined in this column. See Appendix A for explanation of dataset, variable and estimator acronyms
Table 18.7  Evidence of income effects on child health and related health behaviour

<table>
<thead>
<tr>
<th>Authors</th>
<th>Country &amp; data</th>
<th>Sample</th>
<th>Income measure</th>
<th>Measures of health / health behaviour</th>
<th>Estimator</th>
<th>Effect on health</th>
<th>Effect on health behaviour &amp; medical care</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mocan et al (2013)</td>
<td>US Natal Detail Files &amp; CPS Annual Demographic Files 1989-2004</td>
<td>Singleton births to unmarried mothers &gt;19 yrs.</td>
<td>Weekly earnings instrumented by measure of skilled-biased technology shocks</td>
<td>Birth weight, gestation age, prenatal care, smoking &amp; drinking during pregnancy</td>
<td>Two sample IV</td>
<td>Small ↑ birth weight &amp; gestational age babies of low education (≤ high school) mothers unlikely to be on Medicaid</td>
<td>↑ prenatal care for low educ. mothers not on Medicaid, No effect on smoking or drinking.</td>
<td>Effects are very small. Doubling of income ↑ birth weight by 100 grams and ↑ gestational age by 0.7 weeks. No effects for high education mothers or those likely to be on Medicaid</td>
</tr>
<tr>
<td>Hoynes et al (2012)</td>
<td>US Vital Statistics Natality Data 1983-99</td>
<td>Infants of low income mothers</td>
<td>Change in maternal income due to EITC reform</td>
<td>LBW, smoking, drinking, prenatal care use</td>
<td>DID $1000 → 6.7-10.8% ↓ LBW</td>
<td>$1000 → prenatal care ↑0.65 ppt smoking ↓1.2 ppt drinking ↓1.1 ppt ↑ maternal nutrition. ↑ smoking during pregnancy, ↑% children born to unmarried parents</td>
<td>Transfer paid women in bottom income decile</td>
<td></td>
</tr>
<tr>
<td>Amarante et al (2011)</td>
<td>Uruguay mortality, natality, social sec. registers 2003-2007</td>
<td>All newborns</td>
<td>Unconditional cash transfer = 50-100% prepayment income to low income women</td>
<td>LBW, maternal smoking, nutrition</td>
<td>DID, FE &amp; RDD 15% ↓ LBW (1.5 ppt of baseline 10 ppt)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fernald and Hidrobo (2011)</td>
<td>Ecuador, 2003-06</td>
<td>Children 12-35 months</td>
<td>Undonditional cash transfer =$15 per month (6-10% mean hh old. exp.) to low income women</td>
<td>HAZ score, haemoglobin concentration,Vitamin A &amp; iron supplements</td>
<td>Randomize d experiment</td>
<td>No effects</td>
<td>↑ vitamin A &amp; iron supplements in rural areas</td>
<td>Transfers paid to women in poorest 40%</td>
</tr>
<tr>
<td>Duflo (2000 &amp; 2003)</td>
<td>South Africa SALDRU 1993</td>
<td>children 6-60 months</td>
<td>Woman’s receipt of pension</td>
<td>HAZ &amp; WHZ scores 2SLS (IV= pension eligibility)</td>
<td>↑ HAZ &amp; WHZ of girls by 1.2 std. dev. after two years</td>
<td></td>
<td>No effect on boys. No effect of pension paid to male on either girls or boys.</td>
<td></td>
</tr>
</tbody>
</table>

Notes: See Appendix A for explanation of dataset, variable and estimator acronyms
### Table 18.8. Cross-country evidence on population health-income inequality association

<table>
<thead>
<tr>
<th>Authors</th>
<th>Countries / region &amp; period</th>
<th>Data</th>
<th>Estimator</th>
<th>(partial) Correlation with income inequality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pascual et al. (2005)</td>
<td>12 EU countries; 1994-2001</td>
<td>ECHP and OECD</td>
<td>Linear RE &amp; FE</td>
<td>LE -, U5MR +</td>
</tr>
<tr>
<td>Cantarero et al. (2005)</td>
<td>12 EU countries; 1994-2001</td>
<td>ECHP and OECD</td>
<td>Linear RE &amp; FE</td>
<td>LE -, U5MR +</td>
</tr>
<tr>
<td>Biggs et al. (2010)</td>
<td>22 Latin American countries; 1960-2007</td>
<td>WDID; GTD; WIID; SEDLAC</td>
<td>Linear RE &amp; FE</td>
<td>LE ns, IMR ns</td>
</tr>
<tr>
<td>Avendano (2012)</td>
<td>34 OECD countries 1960-2008</td>
<td>WIID, OECD</td>
<td>Poisson FE</td>
<td>IMR ns</td>
</tr>
<tr>
<td>Tacke &amp; Waldmann (2013)</td>
<td>93 countries 1999-2005</td>
<td>WIID; WDID; GHN</td>
<td>OLS</td>
<td>LE -, IMR +, U5MR +</td>
</tr>
</tbody>
</table>

Notes: LE – indicates that life expectancy is negatively correlated with income inequality. IMR + indicates the infant mortality is positively correlated with income inequality. Acronyms of other health indicators, datasets and estimators are explained in Appendix Table A2. ns indicates no significant association.

### Table 18.9. US cross region evidence of population health-income inequality association

<table>
<thead>
<tr>
<th>Authors</th>
<th>Regional unit &amp; period</th>
<th>Data</th>
<th>Estimator</th>
<th>Health measure</th>
<th>(partial) Correlation with income inequality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ash and Robinson (2009)</td>
<td>287 MSA 1990</td>
<td>CMF, STF</td>
<td>WLS</td>
<td>Age-adjusted mortality ratio</td>
<td>Varies with size of MSA</td>
</tr>
<tr>
<td>Deaton and Lubotsky (2009)</td>
<td>287 MSA 1980, 1990</td>
<td>CMF; PUMS</td>
<td>WLS</td>
<td>Age-adjusted mortality ratio</td>
<td>No robust association (conditional on racial composition)</td>
</tr>
<tr>
<td>Yang et al. (2012)</td>
<td>3072 counties 1998-2002</td>
<td>CMF</td>
<td>QR</td>
<td>Age-adjusted mortality ratio</td>
<td>+ with effect increasing in magnitude until 80th percentile of mortality</td>
</tr>
</tbody>
</table>

Note: -/+ indicates negative/positive association of health indicator with income inequality. See Appendix Table A2 for explanation of acronyms. MSA is Metropolitan Statistical Area.
<table>
<thead>
<tr>
<th>Authors</th>
<th>Country, region, period</th>
<th>Data</th>
<th>Estimator</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Theodossiou and Zangeldis (2009)</td>
<td>6 EU countries 2004</td>
<td>SOCIOLD</td>
<td>Linear IV</td>
<td>better ADL: – AL + RD</td>
</tr>
<tr>
<td>d'Hombres et al. (2010)</td>
<td>8 former Soviet countries, 2001</td>
<td>LLH</td>
<td>Probit, OLS, GMM</td>
<td>better SAH: +AL, ns RD</td>
</tr>
<tr>
<td>Karlsson et al. (2010)</td>
<td>21 low/middle/high income countries</td>
<td>FORS, WIID</td>
<td>Pooled ordered probit</td>
<td>better SAH: + AI, + RI, –INEQ in rich countries (= better ADL)</td>
</tr>
<tr>
<td>van Groezen et al. (2011)</td>
<td>10 EU 2004</td>
<td>SHARE</td>
<td>OLS</td>
<td>(= mental health)</td>
</tr>
<tr>
<td>Gröningqvist et al. (2012)</td>
<td>Sweden municipalities 1987-2004</td>
<td>Hospital admissions register CNEF</td>
<td>Linear FE</td>
<td>Hospital admission: ns INEQ; (= sickness leave and mortality)</td>
</tr>
<tr>
<td>Nilsson and Bergh (2013)</td>
<td>Zambia 155 constituencies, 72 districts, 9 provinces 2004</td>
<td>LCMS IV</td>
<td>OLS &amp; 2SLS</td>
<td>HAZ: +AL, – RI (constituency ref.), + RI (provincial ref.), +INEQ</td>
</tr>
</tbody>
</table>

Notes: AI=absolute income, RI=relative income, INEQ=income inequality, RD=relative deprivation, RP=relative position, SC=social capital. Read ‘XXX: AI+’ as the health indicator XXX is significantly positively associated with absolute income. Same for RI, INEQ, RD, RP and SC. Similarly: - indicates negative association; ns indicates no significant association; ~AI indicates no consistent evidence in favour of or against AI hypothesis (respectively for INEQ, RI, RD, RP and SC hypotheses); \( \cap \) INEQ indicates an inverse U shape relationship with income inequality. When indicated in second column, regional unit indicates the level at which income inequality and references for relativities are defined. When not indicated inequality / relativities is at national level. See Appendix Table A2 for explanation of other acronyms.
Appendix

Table A1: Descriptions and means of variables used in analyses in section 2

<table>
<thead>
<tr>
<th>Description</th>
<th>US</th>
<th>Netherlands</th>
<th>China</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-assessed health (SAH)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>poor*</td>
<td>1 if self-assessed health is poor, 0 otherwise</td>
<td>0.031</td>
<td>0.015</td>
</tr>
<tr>
<td>moderate/fair</td>
<td>1 if self-assessed health is moderate/fair, 0 otherwise</td>
<td>0.123</td>
<td>0.156</td>
</tr>
<tr>
<td>good</td>
<td>1 if self-assessed health is good, 0 otherwise</td>
<td>0.339</td>
<td>0.222</td>
</tr>
<tr>
<td>very good</td>
<td>1 if self-assessed health is very good, 0 otherwise</td>
<td>0.394</td>
<td>0.168</td>
</tr>
<tr>
<td>excellent</td>
<td>1 if self-assessed health is excellent, 0 otherwise</td>
<td>0.112</td>
<td>0.039</td>
</tr>
<tr>
<td>ln(income)</td>
<td>natural logarithm of equivalent gross household income</td>
<td>10.142</td>
<td>7.786</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>low*</td>
<td>1 if upper secondary education or less, 0 otherwise</td>
<td>0.271</td>
<td>0.413</td>
</tr>
<tr>
<td>middle</td>
<td>1 if post-secondary non-tertiary education, 0 otherwise</td>
<td>0.341</td>
<td>0.245</td>
</tr>
<tr>
<td>high</td>
<td>1 if tertiary education, 0 otherwise</td>
<td>0.388</td>
<td>0.342</td>
</tr>
<tr>
<td>Gender: male</td>
<td>1 if male, 0 if female</td>
<td>0.403</td>
<td>0.470</td>
</tr>
<tr>
<td>Ethnicity: main group</td>
<td>1 for largest ethnic group</td>
<td>0.737</td>
<td>0.879</td>
</tr>
<tr>
<td>Age (years)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20-29*</td>
<td>1 if age is 20-29 years, 0 otherwise</td>
<td>0.183</td>
<td>0.099</td>
</tr>
<tr>
<td>30-39</td>
<td>1 if age is 30-39, 0 otherwise</td>
<td>0.197</td>
<td>0.148</td>
</tr>
<tr>
<td>40-49</td>
<td>1 if age is 40-49, 0 otherwise</td>
<td>0.203</td>
<td>0.188</td>
</tr>
<tr>
<td>50-59</td>
<td>1 if age is 50-59, 0 otherwise</td>
<td>0.217</td>
<td>0.211</td>
</tr>
<tr>
<td>60-69</td>
<td>1 if age is 60-69, 0 otherwise</td>
<td>0.135</td>
<td>0.218</td>
</tr>
<tr>
<td>70+</td>
<td>1 if age is 70+, 0 otherwise</td>
<td>0.066</td>
<td>0.136</td>
</tr>
<tr>
<td>Employment status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>employed*</td>
<td>1 if employed, 0 otherwise</td>
<td>0.564</td>
<td>0.548</td>
</tr>
<tr>
<td>unemployed</td>
<td>1 if not working and report being unemployed, 0 otherwise</td>
<td>0.107</td>
<td>0.028</td>
</tr>
<tr>
<td>disabled</td>
<td>1 if not working and report being disabled, 0 otherwise</td>
<td>0.070</td>
<td>0.044</td>
</tr>
<tr>
<td>retired</td>
<td>1 if not working and report being retired, 0 otherwise</td>
<td>0.134</td>
<td>0.226</td>
</tr>
<tr>
<td>not working</td>
<td>1 if not working and do not report being unemployed/disabled/retired</td>
<td>0.124</td>
<td>0.154</td>
</tr>
<tr>
<td>Number of obs</td>
<td>5,050</td>
<td>4,137</td>
<td>7,694</td>
</tr>
</tbody>
</table>

Notes:
* reference category in the least squares and interval regressions in tables 18.1 and 18.2.
† In US and Dutch surveys respondents report their health in general as being excellent, very good, good, fair (US)/moderate (NL), or poor. In Chinese survey respondents report their health relative to others of their own age as very good, good, fair or poor.
jection of equivalent gross household income in national currency.
‡ Education has been classified using the International Standard Classification of Education (ISCED) for the US and the Netherlands where low education refers to ISCED<4, middle education to ISCED=4, and higher education to ISCED>4 (Unesco Institute for Statistics). For China, low education refers to primary or junior high school degree, middle education to a senior high school degree, and higher education to vocational higher education and university higher education.
Table A2: Acronyms used in Tables

<table>
<thead>
<tr>
<th>Name / definition</th>
<th>Name / definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>BHPS British Household Panel Study</td>
<td>BRFSS Behavioural risk factor surveillance system</td>
</tr>
<tr>
<td>BRFSS Behavioural risk factor surveillance system</td>
<td>CERRA Leiden University Center for Research on Retirement and Aging panel</td>
</tr>
<tr>
<td>CHS US National Centre for Health Statistics</td>
<td>CIDJ Chinese Inequality and Distributive Justice survey project</td>
</tr>
<tr>
<td>CMF Compressed Mortality File of the National Centre for Health Statistics</td>
<td>CNEF Cross-national equivalent file</td>
</tr>
<tr>
<td>ECHP European Community Household Panel</td>
<td>FORS Future of Retirement Survey</td>
</tr>
<tr>
<td>FSUH Financial Survey of Urban Housing</td>
<td>GHN Globalization-Health Nexus database</td>
</tr>
<tr>
<td>GHS General Household Survey</td>
<td>GSOEP German Socioeconomic Panel</td>
</tr>
<tr>
<td>GSN Globalization-Health Nexus database</td>
<td>GTD WHO Global Tuberculosis Database</td>
</tr>
<tr>
<td>HRS Health and Retirement Study</td>
<td>HSE Health Survey of England</td>
</tr>
<tr>
<td>IOT International Obesity Taskforce</td>
<td>LCMS Living Condition Monitoring Study</td>
</tr>
<tr>
<td>LLH Living Conditions, Lifestyle and Health survey</td>
<td>NCD Swedish National Cause of Death Statistics</td>
</tr>
<tr>
<td>NCD Swedish National Cause of Death Statistics</td>
<td>NDH National and nutrition examination survey</td>
</tr>
<tr>
<td>NTS Swedish National Income Tax Statistics</td>
<td>OECD OECD Health Data</td>
</tr>
<tr>
<td>PUMS US Census Public Use Micro Sample</td>
<td>PSID Panel Study of Income Dynamics</td>
</tr>
<tr>
<td>RHS Retirement History Study</td>
<td>SALDRU South African Labour &amp; Development Research Unit survey</td>
</tr>
<tr>
<td>SEDLAC Socio-economic database for Latin America and the Caribbean</td>
<td>SHARE Survey of health, ageing and retirement in Europe</td>
</tr>
<tr>
<td>SHARE Survey of health, ageing and retirement in Europe</td>
<td>SOCIOILD Socio-economic &amp; occupational effects on the health inequality of the older workforce</td>
</tr>
<tr>
<td>STF US Census Summary Tape File 3C</td>
<td>Ulf Statistics Sweden’s Survey of Living Conditions</td>
</tr>
<tr>
<td>UNHDR United Nations Development Report</td>
<td>WDID World Bank World Development Indicators</td>
</tr>
<tr>
<td>WHO various databases</td>
<td>WIID WIDER World Income Inequality Database</td>
</tr>
</tbody>
</table>

**Health measures**

| ADL | Activities of Daily Living |
| U5MR | Under-5 mortality rate |
| GHQ | General health questionnaire (psychological health) |
| HAZ | Height-for-age Z score |
| HSCL | Hopkins Symptoms Checklist |
| IMR | Infant mortality |
| LE | Life expectancy |
| Major diagnosis | Cancer, heart disease, lung disease (McClellan, 1998– minor) |
| Minor diagnosis | Hypertension, diabetes, stroke (McClellan, 1998 – major), arthritis, back pain |
| MR | Mortality rate |
| SAH | Self-assessed health |
| SB | Stillbirth rate |
| WHZ | Weight-for-height z-score |
Data sources

We use data from the Chinese Health and Nutrition Survey (CHNS), the Netherlands Longitudinal Internet Studies for the Social Sciences (LISS) and the American Life Panel (ALP). We thank the organisations that collected and provided these data.

The LISS panel data were collected by CentERdata (Tilburg University, The Netherlands) through its MESS project funded by the Netherlands Organization for Scientific Research.

RAND ALP are proprietary capabilities of the RAND Corporation. They were developed by RAND with its own funds and with the support of numerous clients and grantors who have commissioned social-science and economics research and analysis at RAND.

The National Institute of Nutrition and Food Safety, China Center for Disease Control and Prevention, Carolina Population Center, the University of North Carolina at Chapel Hill, the NIH (R01-HD30880, DK056350, and R01-HD38700) and the Fogarty International Center, all contributed to the collection of the CHNS.

We than the Equality Trust Fund for providing the data to reconstruct the first figure in section 5.
References


TWADDLE, A., 1974. The concept of health status. Social Science and Medicine, 8, pp. 29-38.


WILKINSON, R.G., 1996-last update, Unhealthy societies the afflictions of inequality [Homepage of Routledge], [Online].


