

Kjellsson, Gustav; Petrie, Dennis; Van Ourti, Tom

**Working Paper**

## Measuring income-related inequalities in risky health prospects

Tinbergen Institute Discussion Paper, No. TI 2018-007/V

**Provided in Cooperation with:**

Tinbergen Institute, Amsterdam and Rotterdam

*Suggested Citation:* Kjellsson, Gustav; Petrie, Dennis; Van Ourti, Tom (2018) : Measuring income-related inequalities in risky health prospects, Tinbergen Institute Discussion Paper, No. TI 2018-007/V, Tinbergen Institute, Amsterdam and Rotterdam

This Version is available at:

<https://hdl.handle.net/10419/177697>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**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.*

TI 2018-007/V  
Tinbergen Institute Discussion Paper



# Measuring income-related inequalities in risky health prospects

Gustav Kjellsson<sup>1</sup>  
Dennis Petrie<sup>2</sup>  
Tom (T.G.M.) van Ourti<sup>3</sup>

1: Department of Economics, University of Gothenburg

2: Centre of Health Economics, Monash University

3: Erasmus School of Economics, Erasmus University Rotterdam; 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.

Contact: [discussionpapers@tinbergen.nl](mailto:discussionpapers@tinbergen.nl)

More TI discussion papers can be downloaded at the [Tinbergen Site](#)

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam  
Gustav Mahlerplein 117  
1082 MS Amsterdam  
The Netherlands  
Tel.: +31(0)20 598 4580

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

# Measuring income-related inequalities in risky health prospects

Gustav Kjellsson<sup>a,b</sup>, Dennis Petrie<sup>c,\*</sup>, Tom Van Ourti<sup>d,e</sup>

<sup>a</sup>Department of Economics, University of Gothenburg, Gothenburg, Sweden

<sup>b</sup>Centre of Health Economics at University of Gothenburg (CHEGU), Gothenburg, Sweden

<sup>c</sup>Centre for Health Economics, Monash University, Melbourne, VIC, Australia

<sup>d</sup>Erasmus School of Economics, Erasmus University Rotterdam, Rotterdam, the Netherlands

<sup>e</sup>Tinbergen Institute, Erasmus University Rotterdam, Rotterdam, the Netherlands

---

## Abstract

The measurement of health disparities is a key component for the assessment of health systems. One aspect of these disparities – which hitherto has received limited attention – is the risk people face about their future health. This paper integrates risk into the standard inequality measurement which measures the extent to which disparities in realized health are systematically associated with income. It develops a rank dependent inequality index that considers not only inequalities in expected future health but also the dispersion of individuals' future health prospects. It is useful when a social planner wants to account for risk averse preferences in the assessment of income-related health inequalities. The empirical application using Australian longitudinal data highlights that neglecting risk underestimates income-related health inequalities since the poor were not only expected to be in worse health in the future, but also faced greater dispersion in their future health prospects compared to the rich.

*Keywords:* health inequality, risk, concentration index

*JEL:* D63, I10

---

## 1. Introduction

It has long been established that health disparities are systematically related to income (Wagstaff and van Doorslaer, 2000; Cutler et al., 2012; Van Ourti et al., 2014); and health economists have proposed measuring these disparities using the concentration index (Wagstaff et al., 1991; Van Doorslaer and Van Ourti, 2011). The concentration index captures the extent of health differences across individuals ranked by some indicator of socioeconomic status, say income, and is considered the main workhorse of inequality measurement in health economics (O'Donnell et al., 2008, 2016). Several review articles have summarized the numerous studies measuring these inequalities, and have confirmed that the positive association between income rank and health holds across space and time (Wagstaff and van Doorslaer, 2000; Van Doorslaer and Van Ourti, 2011; Deaton, 2013).

To unravel the reasons behind the persistence of income-related health inequalities, there has – during the last 15 years – been a move away from the cross-sectional towards the longitudinal perspective. One strand of the literature has focused on the evolution of cross-sectional estimates of health inequalities over time (Gravelle and Sutton, 2003; Wagstaff et al., 2003). Later work disentangled the evolution of inequalities within a cohort of individuals, and established that future health declines are negatively related to the initial income position of individuals (Jones and Nicolas, 2004; Van Ourti et al., 2009; Allanson et al., 2010; Allanson and Petrie, 2013a; Petrie et al., 2011; Baeten et al., 2013). One aspect – which hitherto has received limited attention – is the risk people face about their future health.<sup>1</sup> The literature on decisions under risk stresses that individuals' valuation (or anticipated utility) of particular

---

\*Corresponding author: 15 Innovation Walk, Monash University, Clayton, Melbourne, Australia 3800; fax: (+61) 3 9905 8344

*Email addresses:* [gustav.kjellsson@economics.gu.se](mailto:gustav.kjellsson@economics.gu.se) (Gustav Kjellsson), [dennis.petrie@monash.edu](mailto:dennis.petrie@monash.edu) (Dennis Petrie), [vanourti@ese.eur.nl](mailto:vanourti@ese.eur.nl) (Tom Van Ourti)

<sup>1</sup>The role of risk (or insecurity) has been studied in the income inequality literature (e.g. Makdissi and Wodon 2003; Cruces et al. 2004; Bossert and D'Ambrosio 2013; Rohde et al. 2014; Makdissi and Wodon 2018), but is intrinsically different from our case which focuses on the bivariate setting of risky future health prospects across income.

lotteries often depend on the risk involved (Quiggin, 1982; Starmer, 2000). The riskiness of a future health distribution might thus impact on the value (or anticipated utility) the individual places on it, and it may therefore be relevant to analyse how risk varies across the population in many instances. For example, risk appears in the intergenerational setting where parental characteristics at birth (including income) predict the variability of a child's health later in life (Case et al., 2002; Currie, 2009). In the empirical part of this paper we study an Australian example where 11 years into the future the dispersion in individuals' risky health prospects varies substantially with current health, income and age, and to a lesser extent with lifestyles, gender, unemployment, and marital status. When a social planner worries that the exposure to risk is unequally distributed across income, risk will constitute an additional source of income-related health inequalities on top of disparities in the realized level of health. This is the paper's focus.

Considering risk is tantamount to arguing that probabilities matter. We have a social planner in mind that is knowledgeable about future health prospects, i.e. the individual-specific distribution of future potential health outcomes. The planner knows - or is able to estimate - the different potential health levels that might emerge, as well as the objective probabilities of each *future* health level faced by every individual. We assume that a social planner who cares about income-related health inequalities not only worries about the association between the *current* income position and expected future health but also about the association between the *current* income position and the level of dispersion in future health prospects. This might be a legitimate social concern when comparing risk averse individuals facing health lotteries.<sup>2</sup> Higher levels of dispersion in future health prospects may also have implications on individuals' current economic decision making, potentially distorting investment decisions due to hedging against the full distribution of potential health outcomes (Picone et al., 1998; Chou et al., 2003). The social planner might want to account for such preferences in his assessment of income-related inequalities.

Let us make the issue of risky future health prospects and income-related health inequalities more explicit with a stylized example. First, consider a hypothetical society that consists of two individuals, one currently rich and the other poor; and where future health can take three different levels: dead (0), fair health (0.5) and perfect health (1). In a world without risk, we know future health for sure and the measurement of disparities boils down to assessing whether future health levels are distributed in favor of the poor or the rich, or are equally distributed across income. Figure 1 provides an illustration: it shows probability density functions (PDF) of future health for the currently poor and rich individual in 4 hypothetical societies. In society 1 and 2 there are no income-related disparities because future health is equally distributed across the income dimension, while disparities are in favor of the rich in society 3 (pro-rich), and in favor of the poor in society 4 (pro-poor).

The case without risk is a useful benchmark. In a riskless society, each individual faces only one potential future health level, but when adding risk to the picture, we need to distinguish the entire distribution of potential health outcomes that the individual might experience in the future, from the actual future health level they eventually realize. For example, an individual today might face the risk to be dead ten years later with a probability  $2/3$  or be in perfect health with probability  $1/3$ ; but ten years later she finds herself in perfect health. We focus on the former ex-ante perspective.

Figure 2 illustrates our approach using similar hypothetical societies as in Figure 1, but with an added risk dimension. Risk means that the future health levels of the currently poor and rich persons are not known for sure. Instead, they face a particular probability of being in each health level (death, fair health, perfect health). For example in society 5a, the poor and rich individual face a  $2/3$  chance to be dead,  $0/3$  chance to be in fair health, and  $1/3$  chance to have perfect future health. The central question of this paper is whether (and to what extent) the risky future health prospects are distributed in favor of the poor or rich individual.

Examining inequality in society 5a is easy because the poor and rich individual face the same risky future health prospects, i.e. expected health and the dispersion of potential future health outcomes are the same. Even though the poor and rich individually face a risky future, the risk is identical, and no disparities exists. Society 5b depicts a similar situation but with lower dispersion in both individuals' risky health prospects. In societies 6, the poor and rich experience the same expected health level, but different levels of dispersion in future health prospects. In 6a, disparities are pro-rich because the richer individual faces a lower level of dispersion, while the opposite pattern occurs in society 6b. The higher overall expected health in society 6b (as compared to 6a) does not matter because

---

<sup>2</sup>Here we have in mind a broad concept of risk aversion which reflects both the shape of the utility function (i.e. concavity over health) and preferences for less risky utility prospects (Segal and Spivak, 1990; Wakker, 1994). That is, individuals do not only care about their expected utility but also the utility distribution (Allais, 1979; Hagen, 1979; Quiggin, 1982; Tversky and Kahneman, 1991, 1992; Rabin, 2000; Starmer, 2000).

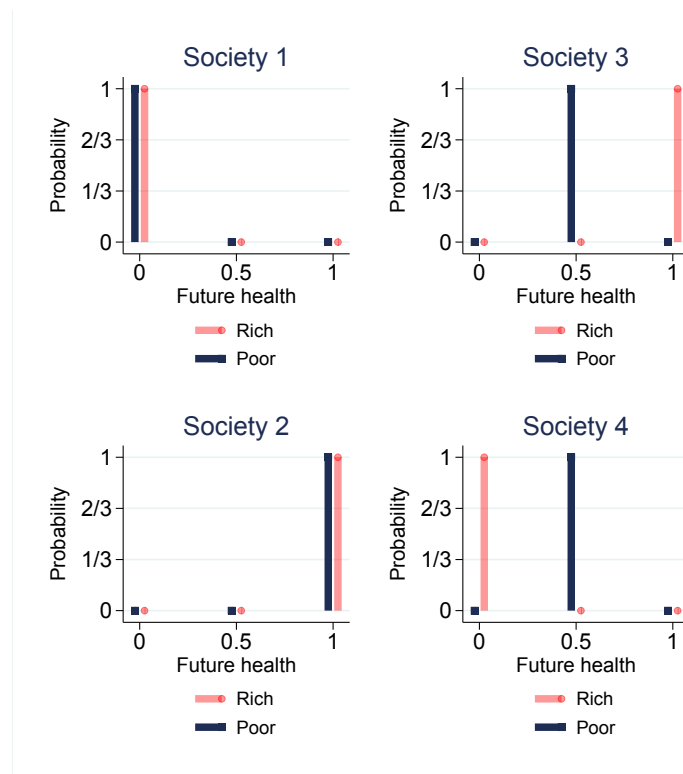


Figure 1: Hypothetical societies without health risk: probability density functions

the poor and rich individual have the same expected health level. In societies 7, expected health differs for the poor and rich individual. Therefore – based on expected health – disparities are in favor of the rich in society 7a and in favor of the poor in society 7b. The dispersion of health prospects has no impact on income-related disparities when we use an absolute measure of dispersion. Indeed, the health prospects of the rich can be obtained by adding (7a) or subtracting (7b) 0.5 units of health from the prospects of the poor. We assume throughout the paper that the social planner considers inequality to be unchanged for absolute changes in all prospects across all individuals, but the framework can be extended to other inequality invariance stances (Kolm, 1976; Erreygers and Van Ourti, 2011; Kjellsson et al., 2015; Allanson and Petrie, 2013c).

This paper makes two important contributions to the literature on health inequality measurement. First, as far as we know, we are the first to introduce a concern about risky health prospects into the standard inequality measurement apparatus used by health economists. Policy makers can use our approach to monitor the extent to which health risk varies with income. We argue that a reasonable assessment of inequalities in ex-ante risky future health prospects will consider inequalities in expected health and inequalities in the dispersion of future health prospects. The new inequality index can be interpreted, and obtained, in two ways. The first estimates the social planner’s valuation of each individual’s risky future health prospects. This valuation is increasing in expected health and decreasing in the dispersion of future health prospects faced by that individual. Next, it evaluates the extent to which these individual-specific valuations are unequally distributed across people with different current income ranks. In other words, the first approach reflects current-income-related inequalities in the social planner’s valuation of risky future health prospects. The second approach is best illustrated using the quantile functions (inverse distribution functions) of the risky health prospects faced by the currently rich and poor. Figure 3 repeats the hypothetical societies in Figure 2 using quantile functions rather than PDF’s. Instead of first aggregating across health quantiles within individuals to obtain the valuation and next across individuals with different income ranks, the alternative approach first estimates current income-related inequalities for every health quantile, and next aggregates the quantile-specific inequalities across health quantiles (i.e. first across income ranks and next across health quantiles). For example, the second

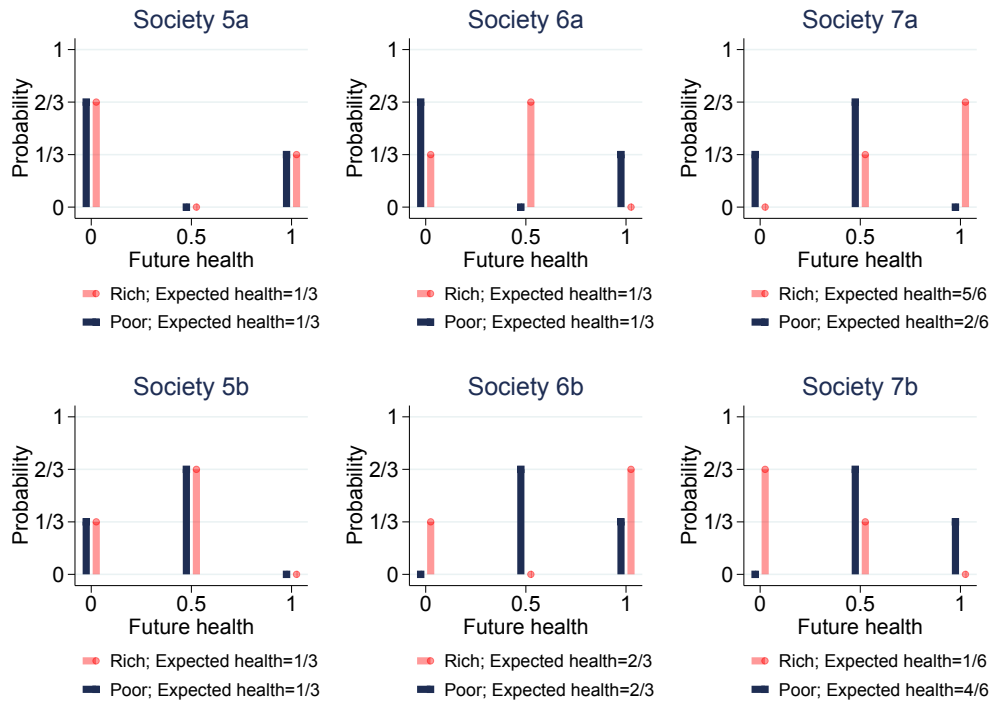


Figure 2: Hypothetical societies with health risk: probability density functions

health tertile of the currently poor in society 6a equals 0, while that of the currently rich individual equals 0.5; hence inequalities in the second health tertile are in favor of the rich. At tertile 3, inequalities are in favor of the poor, while there are no income-related disparities in the first tertile. The overall conclusion about the extent of disparities in risky future health prospects then depends on the importance given to the separate tertile-specific disparities. In the context of risk aversion, more importance is given to the lower tertiles; and hence society 6a has pro-rich disparities in risk. A similar reasoning can be applied to the other societies in Figure 3, and the resulting estimates of income-related disparities will be identical to those obtained from the first estimation procedure.

The second contribution of this paper lies in combining our newly developed measurement apparatus with a covariate decomposition framework (Wagstaff et al., 2003; Allanson and Petrie, 2013b) and applying it to data from the Australian Household, Income and Labour Dynamics in Australia (HILDA) Survey. We find that neglecting risk substantially underestimates income-related health inequalities which highlights the importance of considering the additional inequalities that arise from dispersion. Our empirical results illustrate that the poor were not only expected to have worse health than the rich in the future, but the dispersion of future health prospects were also unequally distributed. Our covariate decomposition shows that the major contributing factors of the greater dispersion in the future health prospects of the poor is not only their older age but also their worse initial health, lower income and to a smaller extent their higher smoking rates.

The remainder of this paper is organised as follows. The methods of inequality measurement of risky future health prospects are laid out in sections 2 and 3, while section 4 outlines the covariate decomposition. The empirical illustration is discussed in section 5 and section 6 concludes.

## 2. Measuring income-related inequalities without risk

While the standard approach to measuring income-related health inequalities cannot deal with risk, it can cope with the case where risk has been removed. To be clear, let us suppose that the social planner, at a given time  $t = 0$ , observes

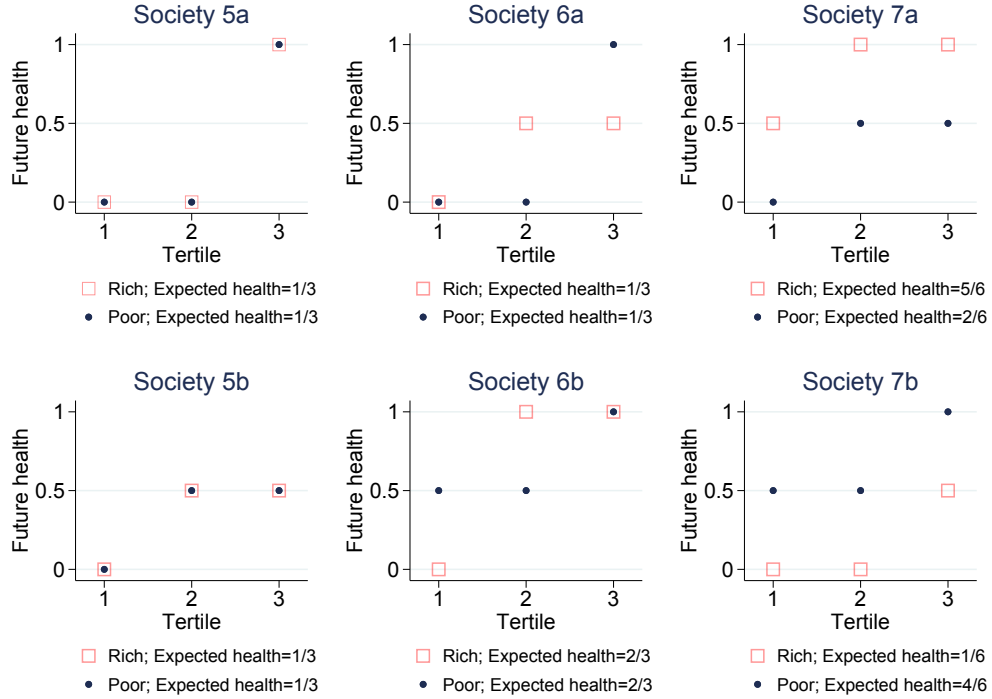


Figure 3: Hypothetical societies with health risk: quantile functions

the health realizations  $H_t^R$  in future period  $t > 0$  ( $t$ -subscripts suppressed hereafter for notational simplicity), and is interested in how these individual realized health outcomes  $h_{it}^R$  relate to individual's income ranks in the initial period  $t = 0$ . We further assume that the individuals are ordered according to their income in the initial period  $Y_0$  (0-subscript referring to initial period suppressed hereafter), such that  $y_1 \leq y_2 \leq \dots \leq y_n$ .<sup>3</sup>

To measure income-related inequalities we focus on rank-dependent indices that impose the absolute inequality invariance criterion – i.e. inequality remaining unchanged when all  $h_i^R$  change by the same absolute amount, but our framework can be extended to other inequality invariance stances (Kolm, 1976; Erreygers and Van Ourti, 2011; Kjellsson et al., 2015). An absolute rank-dependent index  $I(H^R)$  can be written as a weighted sum of health realizations  $h_i^R$  in future period  $t$  where the weights  $z_i$  depend on individual's income ranks  $i = 1, 2, \dots, n$  in the initial period  $t = 0$ :

$$I(H^R, Y) = I(H^R) = \frac{1}{n} \sum_{i=1}^n z_i h_i^R \quad (1)$$

where dependence of  $I(\cdot)$  on  $Y$  is suppressed for notational simplicity further on. The welfare properties of the class of indices in equation (1) have been discussed extensively in the literature (Mehran, 1976; Wagstaff et al., 1991; Bleichrodt and van Doorslaer, 2006; Erreygers and Van Ourti, 2011). Here, we only review the main normative assumptions. First, the weights  $z_i$  sum to zero such that health inequality is zero when everyone has the same health realization  $h_i^R$ .<sup>4</sup> Second, the weights are not decreasing with income in the initial period such that a concentration of health realizations among the poor leads to negative values of the index (pro-poor) and positive values indicate pro-rich inequalities. Third, the class of indices satisfies the principle of income-related health transfers (Bleichrodt and van Doorslaer, 2006) which says that health inequality decreases after a health-transfer from a richer to a poorer

<sup>3</sup>The value of  $I(H^R)$  will depend on the scaling of  $h_i^R$ , but the inequality ordering will remain unaffected as long as health is measured in the same units. This holds for both unbounded and bounded health variables (Erreygers and Van Ourti, 2011; O'Donnell et al., 2016).

<sup>4</sup>The pivotal individual – the individual that delineates the “rich” from the “poor” – will have  $z_i = 0$ .



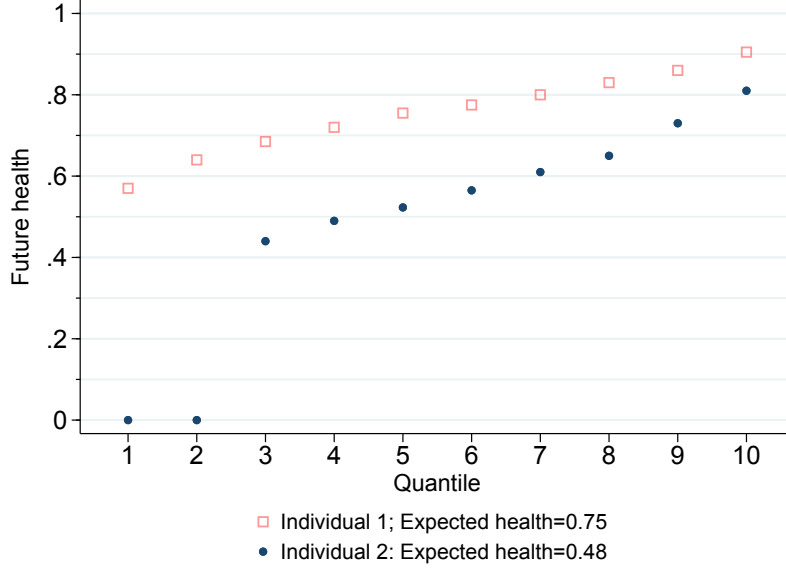


Figure 4: Quantile function of future health prospects of two individuals with  $Q = 10$  quantiles

person. Finally, we do not impose further restrictions on the income rank weights, but in the empirical illustration in section 5 we use weights that are linear in income rank,  $z_i = (2i - 1 - n)/n$ , such that the absolute (or generalized) concentration index emerges.

### 3. Introducing risk into the measurement apparatus

This section introduces risk into the picture. We start by discussing the social planner's valuation of an individual's distribution of future potential outcomes and explore its properties in the first subsection. The second subsection introduces the new index of income-related inequalities and examines the implicit value judgments imposed when replacing the deterministic riskless health values in equation (1) by the social planner's valuations of the risky future health prospects faced by each individual.

#### 3.1. Valuing risky future health prospects: "aggregating across health quantiles"

When individuals face risk about their future health prospects, one might replace the deterministic values  $h_i^R$  in equation (1) by the social planner's valuation of each individual's future health prospects. In order to derive this valuation, we assume that the social planner can observe  $Q$  quantiles of each individual-specific distribution of potential future health outcomes, where the  $q$ th quantile of individual  $i$ 's health distribution in period  $t$  is denoted  $h_i^q$ , so that  $h_i^1 \leq h_i^2 \leq \dots \leq h_i^Q$  for  $q = 1, 2, \dots, Q$ . As an illustration, Figure 4 plots the quantile function (for 10 quantiles) of future health prospects of two individuals where the health variable ranges from 0 to 1. Individual 2 is worse off than individual 1 in terms of expected health ( $0.48 < 0.75$ ). In addition, her distribution is also more dispersed, with particularly more mass at the lower end. If individuals prefer less risky future health prospects, individual 2 would (compared to individual 1) face the double burden of a lower expected future health and a more dispersed distribution.

We measure a social planner's valuation  $V_i$  of the distribution of potential future health outcomes faced by individual  $i$  in period  $t$  using a rank-dependent type expected utility function (see Quiggin, 1982) which is a weighted average of the  $Q$  future health quantiles

$$V_i = \frac{1}{Q} \sum_{q=1}^Q w_q h_i^q \quad (2)$$

We assume that the social planner uses the same valuation function for every individual. The extent to which the social planner incorporates that individuals appreciate health, and are in general risk-averse are determined by the  $w_q$  quantile weights chosen. In order to ensure that the valuation function is strictly increasing in each health quantile, we impose that all quantile weights are strictly positive  $w_q > 0$  for  $q = 1, \dots, Q$ . When the quantile weights equal 1 for all quantiles,  $w_q = w_{q'} = 1$  for any  $q, q'$ , the social planner only considers expected health and does not consider the future health distribution's dispersion. This also suggests that having the quantile weights sum to 1,  $(1/Q) \sum_{q=1}^Q w_q = 1$ , is a natural normalization because the value placed on certain future health is that health level itself.

Next, to guarantee that the social planner incorporates that individuals are in general risk-averse, we impose quantile weights that decrease with  $q$ , i.e.  $w_{q'} \geq w_q > 0$  for  $q' < q$ . Indeed, if the planner considers risk averseness, an improvement of a future health quantile in the lower part of the distribution (a worse health prospect) should improve the valuation function in equation (2) more than an improvement in the upper part of the distribution (a better health prospect). That is because improvements in the lower part of the future health distribution improve expected future health and decrease the dispersion of the future health quantiles, while health improvements in the upper part increase both expected future health and the dispersion.<sup>5</sup> This intuition can be derived from rewriting equation (2) by adding and subtracting  $\bar{h}_i = (1/Q) \sum_{q=1}^Q h_i^q$  to and from  $h_i^q$ :

$$V_i = \frac{1}{Q} \sum_{q=1}^Q w_q (h_i^q + \bar{h}_i - \bar{h}_i) = \bar{h}_i - \frac{1}{Q} \sum_{q=1}^Q w_q (\bar{h}_i - h_i^q) = \bar{h}_i - \frac{1}{Q} \sum_{q=1}^Q h_i^q (1 - w_q) = \bar{h}_i - D_i \quad (3)$$

which shows that the social planner's valuation increases with expected future health  $-\bar{h}_i$  – and where higher dispersion of the future health quantiles  $-D_i = (1/Q) \sum_{q=1}^Q h_i^q (1 - w_q)$  – represents a loss. Equation (3) further shows that an improvement of the  $q$ th health quantile with  $b$  units and a coinciding reduction of the  $q'$ th quantile with  $b$  units leaves expected health unaffected, but reduces the dispersion provided  $q < q'$ .<sup>6</sup>

To keep the discussion general we do not further restrict the quantile weights. However, for the empirical illustration in section 5, we impose a linear function of  $q$ , i.e.  $w_q = (2Q - 2q + 1)/Q$ , such that  $D_i$  equals the absolute (generalized) Gini index; which implies that the social planner's valuation of an individual's future health prospects is obtained by subtracting the absolute Gini of future health prospects from the expected future health.

### 3.2. Income-related inequality in the valuation of risky future health prospects: “aggregating across income ranks” after “aggregating across health quantiles”

Our social planner worries that individuals' risky future health prospects are unequally distributed across current income ranks. The inequality index reflecting this social concern is obtained by replacing the deterministic health realizations  $h_i^R$  in equation (1) by the social planner's valuation function  $V_i$  in equation (2). The resulting index equals a weighted sum of all future health quantiles across all individuals:

$$I(V) = \frac{1}{n} \sum_{i=1}^n z_i V_i = \frac{1}{nQ} \sum_{i=1}^n \sum_{q=1}^Q z_i w_q h_i^q \quad (4)$$

<sup>5</sup>The actual change in the valuation function induced by a health improvement depends on the exact quantile involved and the functional form of  $w_q$ , but an improvement in a higher quantile will, ceteris paribus, lead to a smaller change than an improvement in a lower quantile. Assume that the addition of  $b$  units to health quantile  $q'$  moves it up in the future health distribution such that it becomes larger than what was health quantile  $s$  before adding  $b$  units to health quantile  $q'$  (but smaller than what was initially health quantile  $s + 1$ ). This implies that health quantiles 1 to  $q' - 1$ , and  $s + 1$  to  $Q$  (and their associated weights) are unaffected. Health quantiles  $q' + 1$  to  $s$  neither change, but their associated weights change because  $h_i^s < h_i^{q'} + b$ . In other words, we start from  $h_i^1 \leq \dots \leq h_i^{q'-1} \leq h_i^{q'} \leq h_i^{q'+1} \leq \dots \leq h_i^{s-1} \leq h_i^s < h_i^{s+1} \leq \dots \leq h_i^Q$  and move to  $h_i^1 \leq \dots \leq h_i^{q'-1} \leq h_i^{q'+1} \leq \dots \leq h_i^s < h_i^{q'} + b < h_i^{s+1} \leq \dots \leq h_i^Q$ , leading to a change in the social planner's valuation equalling  $\Delta V_i = (1/Q) \left[ \sum_{t=q'}^{s-1} w_t (h_i^{t+1} - h_i^t) + w_s (h_i^{q'} + b - h_i^s) \right] > 0$ . This shows that  $\Delta V_i$  will be larger when the addition of  $b$  units applies to a quantile in a lower (versus upper) part of the distribution.

<sup>6</sup>The actual change will depend on the functional form of  $w_q$  and on the values of quantiles  $q, q'$ . An exact expression can be derived analogously to that in footnote 5

The index inherits the absolute inequality invariance criterion from the index for the riskless case in equation (1) and therefore will indicate no change in inequality when all health quantiles of all individuals change by the same amount. In addition, the index will indicate no income-related inequalities when all individuals face the same risky future health prospects because the assumptions imposed on the income and health weights guarantee that  $\sum_{i=1}^n \sum_{q=1}^Q z_i w_q = 0$ .

We can use equation (3) to further decompose equation (4) into income-related inequality in expected future health and income-related inequalities due to the dispersion of the future health quantiles

$$I(V) = \left\{ \frac{1}{n} \sum_{i=1}^n z_i \bar{h}_i \right\} - \left\{ \frac{1}{n} \sum_{i=1}^n z_i \left[ \frac{1}{Q} \sum_{q=1}^Q h_i^q (1 - w_q) \right] \right\} = I(\bar{H}) - I(D) \quad (5)$$

which clearly shows that the income weights matter for inequalities in expected future health and dispersion, while the health weights only matter for inequalities due to dispersion. It also shows that pro-rich income-related inequalities in the dispersion of the health quantiles act as a loss while the index is increasing in pro-rich inequalities in expected health. In other words, two societies that face the same association between current income ranks and expected future health, will be evaluated differently by  $I(V)$  if they face a different association between current income ranks and the dispersion of individuals' health quantiles. The society where the poor face more dispersion about their future health outcomes, will be evaluated as being more pro-rich, and vice versa.

The functional form of the weights  $z_i$  and  $w_q$  affects the bounds of the overall index in equation (4), as well as the bounds of the two components in equation (5), and thus determines the relative emphasis on inequalities in expected health  $I(\bar{H})$  versus inequalities in the risk of the future health prospects  $I(D)$ . Since these bounds depend not only on the particular forms chosen for  $z_i$  and  $w_q$ , but also on the type of health variable (i.e. whether it is bounded or not), we derive the bounds only for our particular empirical application in section 5.2.3.

### 3.3. Inequality impact of health quantile changes

The behavior of  $I(V)$  is fully determined by the absolute invariance criterion, and by the assumptions imposed on the income and health weights (income rank weights increase with income and sum to zero; and health quantile weights decrease with  $q$ , are strictly positive and average to one). This section discusses the implications of these assumptions, and in particular how the index reacts to (a) one individual experiencing a change in one particular health quantile; and (b) transfers between different health quantiles of different individuals.

#### *Changing health quantiles of one individual*

Let us start with the simple case where only one individual experiences an increase in one quantile of her future health distribution. In this case, the sign of the income weight – i.e. whether the individual is relatively poor or rich – determines whether the inequality change is pro-poor or pro-rich, with more positive (negative) weights leading to larger inequality changes in favor of the rich (poor). Instead, the health quantile weights only matter for the absolute magnitude of the inequality change, and not for its sign. That is, since all health weights are strictly positive, the sign of  $z_i$  determines the direction of the change of  $I(V)$ ; while improvements of bad future health outcomes will have a bigger impact on the absolute value of the inequality change than further improving already good health outcomes of the same individual. The same reasoning also explains why the most negative weight  $z_i w_q$  is assigned to the poorest ranked individual's worse health outcome, and the most overall positive weight to the richest person's worst health outcome.

#### *Transfers between different health quantiles of different individuals*

Next, we consider the inequality-impact of mean-preserving transfers either between the health quantiles of the same individual, or between the future health distributions of two individuals that may have different income ranks. We base our discussion on the simple case where the transfers are sufficiently small such that they do not affect the associated weights of the other quantiles. Under this assumption, the individuals ( $d$  denoting donor and  $r$  denoting recipient) are confronted with  $h_r^1 \leq \dots \leq h_r^{q^r-1} \leq h_r^{q^r} + b \leq h_r^{q^r+1} \leq \dots \leq h_r^Q$  and  $h_d^1 \leq \dots \leq h_d^{q^d-1} \leq h_d^{q^d} - b \leq h_d^{q^d+1} \leq \dots \leq h_d^Q$  after transferring  $b$  units of health from quantile  $q^d$  of individual  $d$  to quantile  $q^r$  of individual  $r$ . This leads to the following inequality change

$$\Delta I = \frac{1}{nQ} \left\{ \left[ \sum_{i \neq r,d} \left( \sum_{q=1}^Q z_i w_q h_i^q \right) \right] + z_r \left[ \sum_{q \neq q^r} w_q h_r^q + w_{q^r} \left( h_r^{q^r} + b \right) \right] + z_d \left[ \sum_{q \neq q^d} w_q h_d^q + w_{q^d} \left( h_d^{q^d} - b \right) \right] \right\} \quad (6)$$

$$- \frac{1}{nQ} \left( \sum_{i=1}^n \sum_{q=1}^Q z_i w_q h_i^q \right) = \frac{b}{nQ} \left( z_r w_{q^r} - z_d w_{q^d} \right)$$

Equation (6) can be further decomposed into

$$\Delta I = \frac{b}{nQ} \left\{ [z_r - z_d] - \left[ z_r (1 - w_{q^r}) - z_d (1 - w_{q^d}) \right] \right\} \quad (7)$$

showing that the inequality change is partly driven by the impact via the expected health level of both individuals, i.e.  $(b/nQ)(z_r - z_d)$ , and partly by the impact that runs via the dispersion levels, i.e.  $(b/nQ) \left[ z_r (1 - w_{q^r}) - z_d (1 - w_{q^d}) \right]$ . Not unexpectedly, the health weights do not matter for the first part, while the income weights matter for both. It is also intuitive that both equations depend on the magnitude of the transfer ( $b$ ), the income weights ( $z_r$  &  $z_d$ ) and the health weights ( $w_{q^r}$  &  $w_{q^d}$ ) involved. Nevertheless, proposition 1 shows that the inequality change cannot always be unequivocally signed which happens because of the interplay between the income and health weights in equations (6) and (7).

**Proposition 1.** *When a given amount of health is transferred between two different quantiles of two individuals with different levels of income, we can unequivocally sign the inequality change when the income weights of these individuals have an opposite sign. When the income weights have the same sign (positive or negative), the inequality change also depends on the configuration of the two health quantile weights.<sup>7</sup>*

Proposition 1 is best understood by considering all possible configurations of the income and health weights. Table 1 provides such an overview. We consider 8 cases for the income weights. Columns 1 and 2 describe the situation of transfers within an individual or when both individuals have the same income weight. In column 1 both individuals have negative weights while they are positive in column 2. The donating individual is richer in columns 3-5 compared to the receiving individual, and the reverse happens in columns 6-8. In both scenarios we consider 3 cases, i.e. one where both individuals have negative income weights (3 & 6), another where the income weights have opposite signs (4 & 7), and one where both individuals have positive income weights (5 & 8). Finally, we consider three options for the health weights. The first assumes that the health weights are similar, the second that health is transferred from a lower part of the donating individual's future health distribution, towards a higher quantile of the receiving individual; and the reverse happens in the third case.

For transfers between health quantiles of the same individual or between individuals with the same income weight, the sign of the inequality change depends on the configuration of the health weights (and on whether the involved individuals have positive or negative income weights). When the health transfer affects the same health quantiles ( $w_{q^r} = w_{q^d}$ ), inequality remains unchanged<sup>8</sup>; while reducing (increasing) the dispersion of the health quantiles of a poor individual makes inequalities more pro-poor (pro-rich) – column 1 – and the opposite happens when reducing (increasing) the dispersion of the health quantiles of a rich person – column 2. The same intuition can also be obtained from equation (7): since both individuals have the same income weight, income-related inequalities in expected future health will not change, but inequalities in the dispersion of the health quantiles will change by  $z_{r=d}(b/nQ)(w_{q^d} - w_{q^r})$ , which equals 0 when  $w_{q^r} = w_{q^d}$ .

The second scenario (columns 3-5) considers the case where the donating individual is richer than the receiving individual. When the two affected health quantiles are the same for both individuals,  $I(V)$  will unequivocally indicate that income-related health inequalities in future health prospects have become more pro-poor. This happens because

<sup>7</sup>Proposition 1 also holds when the health transfer affects the health weights of the other quantiles (see AppendixA).

<sup>8</sup>When quantiles are the same, one inevitably considers the case of a transfer between two different individuals.

Table 1: Health quantile transfers and the direction of the inequality change  $\Delta I$ 

	same income		richer donates			poorer donates		
	$z_r = z_d$		$z_r < z_d$			$z_r > z_d$		
	negative (1)	positive (2)	negative (3)	opposite (4)	positive (5)	negative (6)	opposite (7)	positive (8)
same quantile $w_{qr} = w_{qd}$	= 0	= 0	< 0	< 0	< 0	> 0	> 0	> 0
lower quantile donates $w_{qr} < w_{qd}$	> 0	< 0	< 0 $\ddagger$	< 0	< 0	> 0	> 0	> 0 $\ddagger$
higher quantile donates $w_{qr} > w_{qd}$	< 0	> 0	< 0	< 0	< 0 $\ddagger$	> 0 $\ddagger$	> 0	> 0

Note:  $\ddagger$ :  $\Delta I > 0$  when the difference between the income weights is small and the difference between the health weights large;  $\ddagger$ :  $\Delta I < 0$  when the difference between the income weights is small and the difference between the health weights large.

$V$  increases/decreases by the same absolute amount for the poorer/richer individual ( $bw_{qr}/Q = bw_{qd}/Q$ ), and since the poorest individual has a smaller income weight ( $z_r < z_d$ ), the inequality change must be pro-poor.

When the health weights differ, we find in most cases that health inequality becomes more pro-poor, but not always (see  $\ddagger$  in columns 3 and 5 of table 1). A pro-rich move in inequality might occur when health is transferred from a rather bad health prospect of a poor person to an already good health prospect of a marginally poorer person ( $z_r < z_d < 0$  and  $w_{qr} < w_{qd}$ ). Or equivalently when the health transfer involves two relatively rich individuals ( $0 < z_r < z_d$ ) and the least rich receives health at a relatively low ranked quantile ( $w_{qr} > w_{qd}$ ). Thus, while a health transfer from rich to poor always induces a pro-poor change in a riskless society, this does not necessarily translate to a situation with risk.<sup>9</sup> This is because transferring a lower (higher) valued unit of health from a rich (poor) donor to a rich (poor), but marginally poorer, receiver who puts higher (lower) value on this unit of health, will make the rich individuals ( $z_r > 0$ ), as a group, better off than the poor ( $z_i < 0$ ). In relation to (7) one may also understand these exceptions as while any health transfer from rich to poor always induces a pro-poor change in inequality in expected health, the effect on inequality in dispersion may be (more) pro-rich. Generally, the sign reversals in columns 3 and 5 of table 1 are more likely the bigger the difference between the health weights and the smaller the difference between the income weights.

The third scenario (columns 6-8), where the donating individual is poorer than the receiving individual, is exactly opposite to the second scenario in columns 3-5. A pro-rich change occurs when the two affected health quantiles are the same; and the change is also mostly pro-rich when the health weights differ. A pro-poor change might occur when a poor person sees a bad health prospect improve at the expense of a better health prospect of a marginally poorer individual (see  $\ddagger$  in column 6), or when a rich person donates from a bad health prospect to a better health prospect of a marginally richer individual (see  $\ddagger$  in column 8).

Proposition 1 and Table 1 indicate that the sign of the inequality impact of a health transfer cannot be unequivocally signed. While this is true, it does not mean that the magnitude of the inequality change shows discontinuities at particular configurations of the income and health weights. Indeed  $(b/nQ)(z_r w_{qr} - z_d w_{qd})$  is a non-linear, but continuous function of the weights, i.e. for a given difference between the health weights and the income weights of the two individuals involved in the transfer, the magnitude of the inequality change is increasing in the income weights and decreasing in the health weights.<sup>10</sup>

<sup>9</sup>Makdissi and Yazbeck (2016) have derived a new class of health inequality indices for the riskless case that relax the principle of income-related health transfers such that rich-to-poor health transfers might lead to pro-rich changes when the poorer individual experiences much better health than the richer individual. While this is reminiscent of the cases with a  $\ddagger$  in columns 3 and 5 of table 1 (and by extension also columns 6 & 8 with  $\ddagger$ ), it is fundamentally different because we consider a situation with risk.

<sup>10</sup>Consider a hypothetical example where  $b$  health units are transferred from individual  $r$  with  $z_r = 0.25$ ;  $w_{qr} = 1.5$  to individual  $d$  with  $z_d = 0.75$ ;  $w_{qd} = 0.5$ . The inequality change then equals  $(b/nQ)0 = 0$ . Changing the configuration of the weights to  $z_r = 0.25$ ;  $w_{qr} = 1.75$ ;

### 3.4. Alternative specification: “aggregating across health quantiles” after “aggregating across income ranks”

In the introduction, we mentioned a second approach to derive the index in equation (4). It consists of first estimating income-related inequalities for every health quantile – i.e. aggregate a particular health quantile across individuals with different *income ranks* –, and next obtaining the final index as a weighted sum of these quantile-specific income-related inequalities. Notably,  $I(V)$  can be rearranged as a weighted mean of the income-related inequalities in each health quantile

$$I(V) = \frac{1}{nQ} \sum_{i=1}^n \sum_{q=1}^Q z_i w_q h_i^q = \frac{1}{Q} \sum_{q=1}^Q w_q \left( \frac{1}{n} \sum_{i=1}^n z_i h_i^q \right) = \frac{1}{Q} \sum_{q=1}^Q w_q I(H^q) \quad (8)$$

Expressing the index in this format explicitly illustrates how inequalities in each quantile are valued differently depending on its location in the distribution. Given the restrictions imposed on  $w_q$ , more weight is given to income-related inequalities at lower health quantiles and lesser weight to inequalities in quantiles in the upper part of the health distributions.

## 4. Decomposing inequality into the contribution of covariates

When developing the inequality measure, we assumed that in the initial period we are able to ex-ante observe the complete distribution of individual  $i$ 's health prospects for some future period. In reality, however, we are only able to observe the ex-post realization of this distribution. The empirical version of our ex-ante approach therefore needs to be evaluated ex-post using the observed realization of similar individuals as approximations of the health distribution individuals' faced (or ex-ante using the relationships observed in ex-post data for extrapolation into the future for individuals alive today). Section 5 describes the details of the estimation approach used to predict the complete distribution of potential health outcomes for each individual in a particular future period  $t > 0$  given the information available in the initial period, but its essence consists of exploiting variation in initial characteristics across individuals to predict variation in future health quantiles.<sup>11</sup> Variation in the estimated quantiles will thus be dependent on the characteristics that are included in the predictive model. It is therefore useful to understand the variation in these initial characteristics which are associated with the inequality in expected future health and the dispersion of future health prospects. For the exposition of the decomposition, it suffices to assume that the  $Q$  quantiles for each individual  $h_i^q$  can be linearly related to the individual's characteristics  $x_i$  in the initial period, i.e.  $t = 0$ .

$$h_i^q = \beta_0^q + \beta_1^q x_{i1} + \beta_2^q x_{i2} + \dots + \beta_K^q x_{iK} = \beta_0^q + \sum_{k=1}^K \beta_k^q x_{ik} \quad (9)$$

with  $\beta_k^q$  measuring the partial association between  $h_i^q$  and the individual characteristic  $x_{ik}$  at quantile  $q = 1, \dots, Q$ . Analogous to the standard concentration index,  $I(V)$  can be decomposed into contributions from a set of covariates following a similar argument as in Wagstaff et al. (2003).<sup>12</sup> Substituting equation (9) into the overall inequality index in equation (4) then yields

$$I(V) = \frac{1}{nQ} \sum_{q=1}^Q \sum_{i=1}^n z_i w_q \left( \beta_0^q + \sum_{k=1}^K \beta_k^q x_{ik} \right) = \sum_{k=1}^K \left[ \left( \frac{1}{Q} \sum_{q=1}^Q w_q \beta_k^q \right) I(x_k) \right] \quad (10)$$

$z_d = 0.75$ ;  $w_{q,d} = 0.75$  results in  $\Delta I = (b/nQ)(-0.125) < 0$ , while  $z_r = 0.25$ ;  $w_{q,r} = 1.25$ ;  $z_d = 0.75$ ;  $w_{q,d} = 0.25$  results in  $\Delta I = (b/nQ)0.125 > 0$ . Similarly, when we fix the health weights to 1.5 and 0.5, but let the income weights change to (0.125, 0.625) and (0.375, 0.875) we obtain respectively  $\Delta I = (b/nQ)(-0.125)$  and  $\Delta I = (b/nQ)0.125$ .

<sup>11</sup>To focus on the illustration of our new measurement apparatus, we have chosen not to use the additional longitudinal data available to further model health dynamics over time. As explained in the discussion section 6, our empirical approach only requires two data points and is therefore also useful when less rich data is available.

<sup>12</sup>The limitations of the decomposition method suggested by Wagstaff et al. (2003) have recently been highlighted by Heckley et al. (2016) and Erreygers and Kessels (2013) among others, including that it is a health-focused decomposition of a two-dimensional index (it does not consider the role the determinants may play in determining income ranks). This critique is less valid in this case as the ex ante approach taken here assumes the income ranks are fixed based on initial income (the initial characteristics predict future health but initial income rank is pre-determined).

which illustrates that the contribution of each covariate is given by the rank dependent inequality index of that covariate,  $I(x_k) = (1/n) \sum_{i=1}^n z_i x_{ik}$ , weighted by a weighted mean of the  $\beta_k^q$ 's at each quantile,  $(1/Q) \sum_{q=1}^Q w_q \beta_k^q$ . This equation can be further decomposed into how much each covariate contributes to the inequality in expected health outcomes and to inequality in dispersion:  $\bar{\beta}_k I(x_k) = (1/Q) \sum_{q=1}^Q \beta_k^q$

$$I(V) = \sum_{k=1}^K [\bar{\beta}_k I(x_k)] - \sum_{k=1}^K \left[ D(\beta_k^q) I(x_k) \right] \quad (11)$$

The expected health part equals the summed inequality indices of the covariates weighted by the mean of their coefficients at each quantile,  $\bar{\beta}_k = (1/Q) \sum_{q=1}^Q \beta_k^q$ , and the dispersion part equals the sum of the inequality indices of the covariates weighted by the dispersion of their coefficients over the quantiles (denoted as  $D(\beta_k^q) = (1/Q) \sum_{q=1}^Q (1 - w_q) \beta_k^q$ ). Alternatively, we can decompose  $I(V)$  into a weighted mean of the sums of the covariate contributions in each quantile, weighted by  $w_q$ :

$$I(V) = \frac{1}{Q} \sum_{q=1}^Q w_q I(h^q) = \frac{1}{Q} \sum_{q=1}^Q w_q \sum_{k=1}^K \beta_k^q I(x_k) \quad (12)$$

## 5. Empirical illustration

### 5.1. Data

We use the Household, Income and Labour Dynamics in Australia (HILDA) annual survey. HILDA is a household-based panel study in Australia, which covers a large national representative sample of Australian households occupying private dwellings and includes a wide range of topics (Summerfield et al., 2014). Wave 1 (2001) consisted of 7,682 households and 19,914 individuals. The current analysis considers the sample who answered the full questionnaire in 2002 and tracks their health until 2013. The sample was matched to the National Death Index to obtain the year of death. The income variable used in this study is the equivalised household income in 2002 which accounts for the number of adults and children in the household using the OECD equivalence scale.<sup>13</sup> We have dropped the 0.5% poorest of the population in 2002 to remove negative incomes as we use the natural log of equivalised household income as an explanatory variable in the subsequent regressions, and symmetrically we also removed the 0.5% richest. Our health measure is derived in terms of Quality Adjusted Life Years (QALYs) from the SF6D and ranges from zero to one, where dead individuals are assigned a value of zero for all future periods.<sup>14</sup>

### 5.2. Empirical method

#### 5.2.1. Estimating future health prospects

Generally, the nature of the health variable has implications on how one may estimate the future health prospects. In our application health ranges from 0 to 1 where the minimum value 0 indicates death, but the minimum value observed conditional on survival is approximately 0.3. Therefore our approach has to both account for the extreme outcome, death, and restrict the possible predicted health quantiles to fall between the bounds of 0.3 and 1 when they are predicted to be alive. In order to capture these two dimensions, we estimate a two-part model (Duan et al., 1983) for each future year 2003-2013 where we include a linear covariate set of initial (2002) conditions  $x_i$ : health (QALYs), equivalised household income (expressed as the natural logarithm of 2013 Australian dollars in thousands), age (and age squared), sex, highest education qualification (categorized as degree or higher, post school certificate, secondary school completion or less than secondary school completion), unemployment status, marital/civil status (categorized as being single; being married or having a partner; or being separated/divorced or a widow(er)), smoking

<sup>13</sup>This scale assigns a weight of 1 to the first adult and 0.5 to all other adults and 0.3 to all children (those aged under 15 years old).

<sup>14</sup>The QALYs are derived from the responses to the SF-36 questionnaire using the SF-6D preference-based algorithm (Brazier et al., 2002). In order to combine morbidity and mortality into a single metric, health economists have elicited the value of particular health states by asking individuals to make decisions trading off quality versus length of life. QALY weights are derived for health states in comparison to full health which is given a QALY of one and death which is given a QALY weight of zero (see Drummond et al. (2015) for more details).

status (categorized as never, current, or former smoker), dummy for being a risky drinker (drink at least on average once per week and have more than 6 standard drinks on a usual occasion), immigrant status (dummies for being non-Australian born with a English or non-English speaking origin). These initial conditions should be interpreted as predictors of an individual's future health distribution, and are not included to obtain the causal effect of  $x_i$  on the health distribution. Therefore, we include potentially endogenous predictors among the  $x_i$ 's as including them still provides unbiased predictions of the future health quantiles for each individual. Table 2 shows the descriptive statistics of these covariates along with the absolute concentration index of each variable ( $AC(x)$ ) with the initial income rank.

Table 2: Descriptives

VARIABLES ( $n=11,053$ )	MEAN	STD. DEV.	$AC(x)$
Health 2002	0.76	0.12	0.015
Health 2003	0.76	0.13	0.012
Health 2008	0.72	0.20	0.019
Health 2013	0.67	0.26	0.028
Equivalised Household Income (\$'000s per annum)	38.70	21.20	0.312
Age (years old)	43.03	17.28	-1.525
Male	47%		0.017
Education			
Less than secondary school completion	40%		-0.072
Secondary school completion	14%		0.001
Post school certificate or diploma	26%		0.010
Degree or higher	19%		0.061
Unemployed	4%		-0.012
Marital Status			
Single	23%		0.001
Married or living with partner	64%		0.028
Divorced/Separated/Widowed	13%		-0.028
Smoking Status			
Never	49%		0.018
Current	24%		-0.012
Former	27%		-0.006
Risky Drinker	7%		0.001
Country of birth			
Australia	78%		0.011
Non-English speaking	12%		-0.015
English speaking	11%		0.0042

(a) Note: Our sample excludes 16.8% (1857) of people who are dropped due to missing health or other explanatory factors in 2002: mostly due to the SF12 not being fully completed (1772 SF36 not completed as this appears in the self completion questionnaire of the survey rather than the telephone component). 44, 382 and 792 of the 11,053 are recorded as dead by 2003, 2008 and 2013 respectively and all individuals are included in all mortality regressions. Note that in 2003, 2008 & 2013 only 8,981 (82%), 7,552 (71%) & 7,354 (72%) of those alive report their health in each year respectively. Only these people who are alive and report their health are included in the quantile regressions though the predicted future health prospects are calculated for everyone based on their initial characteristics. These health averages include zero health for those that are reported as dead.

The first part of the two part model estimates the probability of being alive in future period  $t > 0$ , i.e.  $h_i > 0$ , using a binary model,  $p_i = P(h_i > 0 | x_i) = F\left(\alpha_0 + \sum_{k=1}^K \alpha_k x_{ik}\right)$ . Assuming  $F(\cdot)$  is the CDF of the logistic distribution yields (a standard logit):



$$P(h_i > 0 | x_i) = \frac{1 - \exp\left(\alpha_0 + \sum_{k=1}^K \alpha_k x_{ik}\right)}{\exp\left(\alpha_0 + \sum_{k=1}^K \alpha_k x_{ik}\right)} \quad (13)$$

The second step is influenced by the work of Flores and O'Donnell (2016) on estimating household exposure to the risk of catastrophic medical payment using quantile regressions (e.g. Koenker and Bassett, 1978). We estimate a logistic quantile regression conditional on survival. The logistic quantile regression is a convenient way to restrict the predicted health quantiles to lie within particular bounds because from an asymptotic perspective quantiles are invariant to monotonic transformations (Bottai et al., 2010; Orsini and Bottai, 2011). In practice, this involves first transforming the health outcome onto a continuous scale between positive and negative infinity<sup>15</sup> using a logistic transformation  $g(h_i) = \ln\left(\frac{h_i - 0.3}{1.001 - h_i}\right)$ , and then use this transformed variable as the dependent variable in the linear quantile regressions (conditional on survival):

$$Q_{g(h_i)}(q | h_i > 0, x_i) = \psi_0^q + \sum_{k=1}^K \psi_k^q x_{ik} \quad (14)$$

The predicted quantiles on this continuous scale between negative and positive infinity are then re-transformed back onto the bounded scale:<sup>16</sup>

$$\hat{h}_i^q = \hat{Q}_h(q | h_i > 0, x_i) = \frac{1.001 \exp\left(\hat{\psi}_0^q + \sum_{k=1}^K \hat{\psi}_k^q x_{ik}\right) + 0.3}{1 + \exp\left(\hat{\psi}_0^q + \sum_{k=1}^K \hat{\psi}_k^q x_{ik}\right)} \quad (15)$$

Asymptotically an infinite number of quantiles can be estimated but in a finite sample there will only be a finite possible number of quantiles with “different” estimated regressions (Portnoy, 1991; Melly, 2005). Also, when estimating the quantiles in the distribution’s tails the uncertainty increases as the statistical power for these regressions are based on the support (size of the samples) on either side of the quantile. We estimate equation (14) for  $q = 0.01, 0.02, \dots, 0.99$ , and use the predictions in equation (15) to provide an estimate of the distribution of possible future health outcomes each individual is expected to face in each future year conditional on survival. To combine mortality and morbidity, these predictions are used jointly with the individuals’ predicted mortality risk from the logit model,  $\hat{P}(\text{mortality}) = 1 - \hat{p}_i = 1 - \left\{ \left[ 1 - \exp\left(\hat{\alpha}_0 + \sum_{k=1}^K \hat{\alpha}_k x_{ik}\right) \right] / \exp\left(\hat{\alpha}_0 + \sum_{k=1}^K \hat{\alpha}_k x_{ik}\right) \right\}$ , to obtain an estimate of the full distribution of possible future health outcomes, including death. We generate an approximation of the predicted future health distribution by drawing 1000 observations for each individual from the predictions of the two parts of the model.<sup>17</sup> By probability  $\hat{P}(\text{mortality}) = 1 - \hat{p}_i$  these draws are assigned the value of 0, and by probability  $\hat{P}(\text{alive}) = \hat{p}_i$  these are drawn from the predictions of the 99 quantiles. These 1000 draws are then ordered such that  $h_i^1 < h_i^2 < \dots < h_i^{1000}$  provide the predicted 1000 quantiles for each individual’s future health prospects.

### 5.2.2. Computing inequality indices

In order to use these estimates of  $\hat{h}_i^q$  (we suppress  $\wedge$  from now on to simplify notation) to compute the valuation of future health prospects for each individual  $V_i = (1/Q) \sum_{q=1}^Q h_i^q - (1/Q) \sum_{q=1}^Q h_i^q (1 - w_q)$ , and the income related inequality in the valuation  $I(V)$ , we need to impose weights for both health quantiles  $w_q$  and income rank  $z_i$ . In this illustration, we use weights linear in health quantiles, defined as  $w_q = (2Q - 2q + 1)/Q$ , such that the second part of the valuation reduces into the absolute Gini coefficient of health quantiles  $G_i(H^q) = (1/Q) \sum_{q=1}^Q h_i^q [1 - ((2Q - 2q + 1)/Q)]$ .

<sup>15</sup>In practice one needs to make the bounds slightly bigger than the true ones otherwise the transformed variable will be undefined as it will be equal to positive or negative infinity.

<sup>16</sup>So the retransformed quantiles are now always  $>0.3$  and  $<1.001$ . One could choose values closer to the actual minimum and maximum (to further “stretch” the tranformed continous variables to negative and positive infinity) but this makes an extremely small difference to the final predicted quantiles.

<sup>17</sup>Using a 1000 draws yields expected conditional probabilities of being dead equal to the mortality rate at a one decimal precision (corresponding to 0.1% at most). Making more than 1000 draws (to further improve the precision) has a very minor impact on our results and increases the computational burden.

In addition, we apply linear income rank weights that correspond to the standard weights for the absolute concentration index, i.e.  $z_i = (2i - 1 - n)/n$ . Thus, the overall inequality index  $I(V)$  we apply equals an absolute concentration index of the valuation of the future health prospects:

$$I(V) = \frac{1}{n} \sum_{i=1}^n V_i \left( \frac{2i - 1 - n}{n} \right) = AC(V)$$

$$AC(V) = AC(\overline{H}) - AC[G(H^q)] \quad (16)$$

where its two components are the absolute concentration index of expected health and the absolute concentration index of dispersion (where dispersion is measured by the absolute Gini coefficient of individual health quantiles). We also consider the alternative decomposition from equation (8)

$$AC(V) = \frac{1}{Q} \sum_{q=1}^Q \left[ \frac{2Q - 2q + 1}{Q} AC(h^q) \right] \quad (17)$$

computing the absolute concentration index for each of the 1000 health quantiles.

### 5.2.3. Bounds of the inequality indices

To interpret the results it is useful to consider the value of the indices when inequality is at its maximum and minimum, as well as the relative importance placed on the two parts of the index. Specifically, for the linear weights applied in our application, the overall index attains its maximum (minimum) value, 0.25 (-0.25) when all health prospects of all individuals above (below) the median income rank have full health in all health quantiles, and individuals below (above) the median rank have no health in any health quantile.<sup>18</sup> The bounds of the inequality index in expected health also equals -0.25 and 0.25 (and it attains its extreme values under the same scenario as the overall index where there is no risk in future health prospects), whereas the inequality index due to dispersion ranges from -0.0625 to 0.0625 (and is equal to its maximum value when the richest 50% of the population have certain future health and the poorest 50% of the population have a 50% chance of full health and 50% chance of zero health).<sup>19</sup> The difference between these bounds is informative for the relative emphasis the index puts on inequality in expected health and inequality of dispersion. Using another set of weights may imply a larger relative emphasis on the second part of the index, although the general properties imposed in section 3 imply that the upper bound of the second part of the index will never exceed the upper bound of the first.

### 5.2.4. Estimating the covariate decomposition

The adjustment of the health quantile estimations using a two-part model to combine morbidity and mortality introduces non-linearities that also affects the decomposition derived in section 4 (introducing any type of non-linearity has a similar effect). A simple solution to this problem is to estimate equation (9) running a linear regression at each quantile of the simulated distribution obtained from the draws of the two-part model in order to get a first order approximation. Using these coefficients and the specific set of linear weights the general decomposition in equation (11) becomes

<sup>18</sup>Having a bounded health variable affects the values the index can attain as a given total amount of health cannot (except for very specific cases) be concentrated in just the richest or poorest individual (Erreygers, 2009), or in the highest health quantile. In this case since the  $w_q$ 's are positive for all quantiles and the  $z_i$ 's are positive for 'rich' individuals (above the median income rank), the maximum value of  $I(V)$  equals  $\left(1/n^2\right) \left(1/Q^2\right) \left[\sum_{i=n/2}^n (2i - 1 - n)\right] \left[\sum_{q=1}^Q (2Q - 2q + 1)\right] = \left(1/n^2 Q^2\right) [2(n/2 + n)(n/2)(1/2) - (1 + n)(n/2)] [(2Q + 1)Q - 2(1 + Q)(Q/2)] = \dots \approx 1/4$ . The same reasoning leads to the minimum value of  $-1/4$ .

<sup>19</sup>The dispersion part will be maximized when each rich individual (income rank above the median) has the same level of health for each of its future health quantiles, i.e.  $G(H^q) = 0$ , and when the poor individuals (income rank below median) have  $G(H^q) = 0.25$ . That is, full health for all quantiles above the median and zero health for all quantiles below.

$$AC(V) = \sum_{k=1}^K AC(x_k) \left[ \frac{1}{Q} \sum_{q=1}^Q \frac{(2Q-2q+1)}{Q} \beta_k^q \right] = \sum_{k=1}^K \bar{\beta}_k AC(x_k) - \sum_{k=1}^K D(\beta_k^q) AC(x_k) \quad (18)$$

where  $D(\beta_k^q) = (1/Q) \sum_{q=1}^Q [(2Q-2q+1)/Q] \beta_k^q$ .<sup>20</sup> Thus, we calculate the contribution to  $AC(V)$  of each covariate as the product of the absolute concentration index of that covariate,  $AC(x_k)$ , and a weighted sum of the coefficients in each quantile  $(1/Q) \sum_{q=1}^Q [(2Q-2q+1)/Q] \beta_k^q$ , and the contributions to the expected health and dispersion parts of  $AC(V)$  are respectively weighted by  $\bar{\beta}_k$  and  $D(\beta_k^q)$ .

### 5.3. Results

#### 5.3.1. Mortality & quantile regressions

A snapshot of the regressions underlying the graphs is provided in AppendixB. We show the logistic regression results for mortality and the estimated coefficients for a sample of the quantile regressions (for quantiles at 5%, 25%, 50%, 75%, 95%) for three future years (2003, 2008, 2013).<sup>21</sup> The results from these regressions are used to simulate the expected health quantiles for each individual (unconditional on survival) for each future year taking into account both mortality and their health quantiles conditional on survival.

#### 5.3.2. “Summing over individuals”

We first look at how the expected future health and the estimated future dispersion for each individual vary by their income rank in 2002. Figure 5a shows the local mean for expected health in selected future years by 2002 income rank, and how the local mean of actual health in 2002 varied by income rank in 2002 for benchmarking purposes. The downward shift of the curves implies that all income groups expect worse health the further into the future they look, and the general shape of the curves implies that the rich in 2002 expect to have better health in the each future year compared to the poor, especially for the poorest 30% of the population. For the same part of the population we see that the slope of the local mean becomes steeper over time, implying that, in general, the rate of health decline is expected to be the greatest for the poorest in 2002. Figure 5b shows the local mean of the dispersion individuals face by 2002 income rank. In general, the poor in 2002 also face greater dispersion in their future health prospects, despite lower future expected health levels, implying that they particularly face less favorable future health levels in the lower quantiles.<sup>22</sup> The level of dispersion also increases for all income groups over time. From 2003 to 2005 and again to 2007 we further see that the slope of the relationship between dispersion and 2002 income rank becomes steeper with the level of dispersion increasing quicker for the poor compared to the rich – again especially for the poorest 30%. This means that the decline in health levels over time was much more evenly spread across the health quantiles for the rich compared to the poor. However as we move further into the future we see a more uniform shift up in the level of dispersion for all income classes.

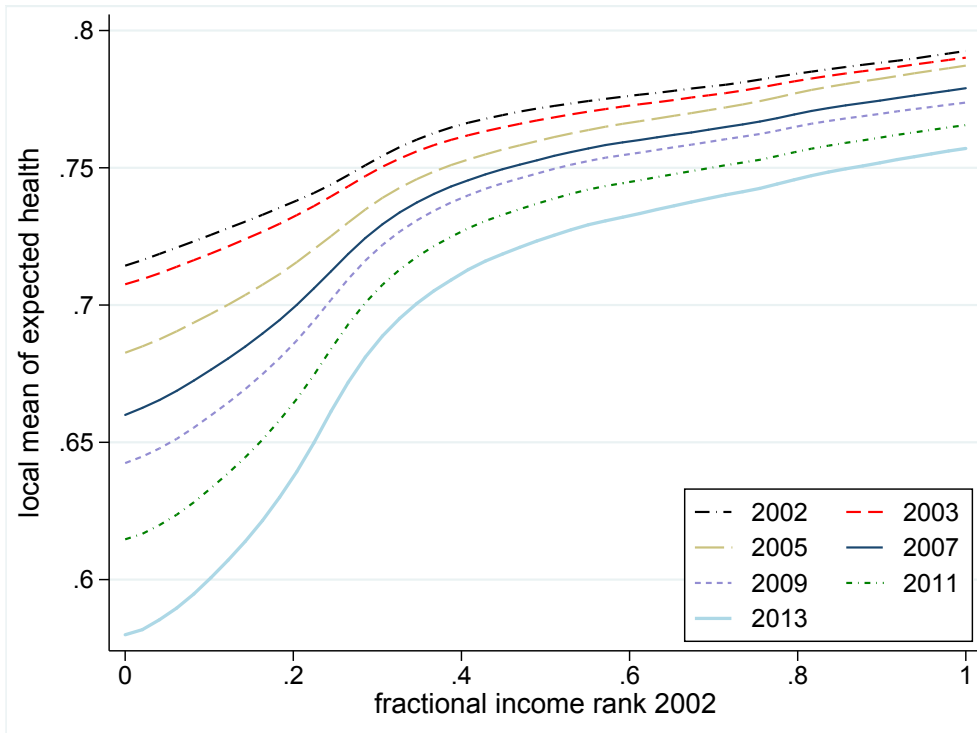
#### 5.3.3. “Summing up the inequality over quantiles”

Figure 6 illustrates the complementary approach to examine the level of income-related inequality by plotting the income-related inequality as measured by the absolute concentration index in each  $q$  of the 1000 predicted quantiles, where 1 is the worse quantile for all individuals and 1000 is the best quantile for all individuals, for selected future

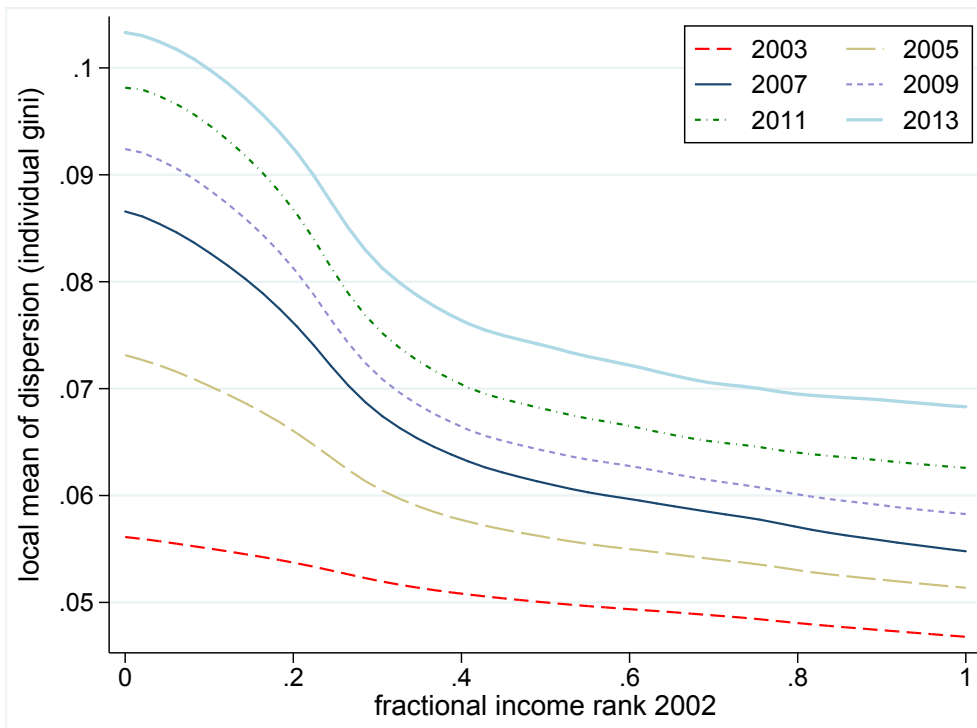
<sup>20</sup>When approximating equation (9) by estimating a linear regression in each quantile, we introduce an error term as in the standard decomposition by Wagstaff et al. (2003). As the health quantiles are predicted using the same covariate set, these errors are close to negligible. The maximum contribution (in absolute terms) of the error from the linear approximation to the total inequality, or any of its two components, in any year is 1.1%. For simplicity, we do not include the error term in equation (18).

<sup>21</sup>Because of the non-linear nature of the two part model and the logistic quantile regressions, the coefficients themselves are largely uninformative of each variable’s marginal impact on future health prospects. Instead, AppendixC provides a graphical presentation of the covariate decomposition results for each quantile according to equation (12) where the contribution is the approximate marginal impact of each (set of) covariate (unconditional on survival) rescaled by  $AC(x_k)$ .

<sup>22</sup>Note that when expected health is very close to the bounds (0 or 1), it follows that dispersion has to be low. This has implications for the bounds of the index and its subcomponents (see section 5.2.3).



(a) Local mean of expected health per year



(b) Local mean of dispersion per year

Figure 5: Local polynomial mean of expected health and dispersion per year

years. The absolute concentration index varies between quantiles reaching a maximum around 0.045 for most years, which is large considering that the upper bound of the index equals 0.25. For the worse quantiles in all future years the absolute concentration index approaches zero indicating a somewhat equal distribution of the worst health outcomes as all income groups face some future chance of dying. As we move up the quantiles the absolute concentration index rapidly increases with the poor on average having increasingly worse potential health outcomes for these lower quantiles compared to the rich. The quantile where the level of income-related inequality is at its maximum slightly increases the further into the future one looks with the maximum quantile being the 175th quantile out of 1000 by the year 2013 - this shift to the right can be explained by an overall increasing level of mortality and morbidity. On the other hand, the shift upwards indicates that the poor experience increasing mortality and health declines (over time) as compared to the rich.<sup>23</sup> For each year, the level of inequality in each quantile falls after the maximum point in a rather linear fashion (with varying slopes) down to a narrow interval between 0.005 and 0.01. Thus, the best possible future health level of each individual are distributed more or less as unequal in 2013 as in 2003. As discussed earlier the weighted sum of the inequality in each quantile, with the greatest weight placed on the worse quantile and the least weight on the best quantile, equals the overall level of inequality.

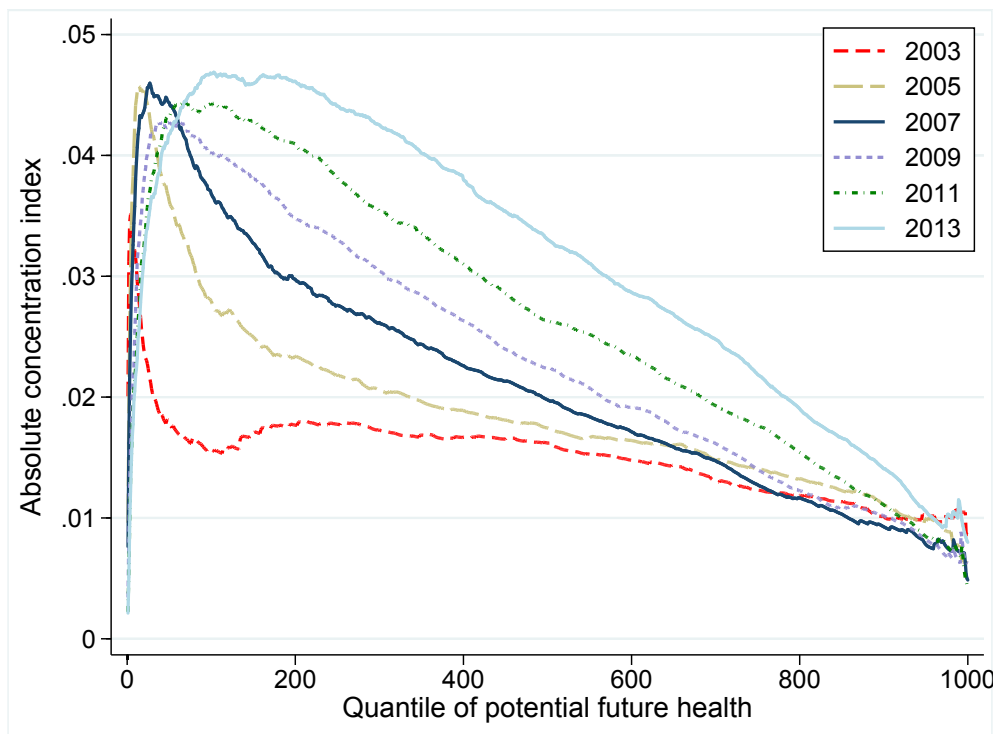


Figure 6: Absolute concentration index per quantile

#### 5.3.4. “Total inequality - decomposed into contributions due to expected health and risk”

Figure 7 outlines for each year the income-related inequality in the total valuation of future health prospects  $V_i$ , as well as broken down in accordance with equation (16) into the inequality in expected health and the inequality in dispersion. For comparison the graph also presents the level of income-related inequality in realized outcomes in 2002. When interpreting the level of inequality one should remember that the upper bound equals 0.25 for the total index, as well as for expected health, but only 0.0625 for the dispersion part. In terms of inequality in expected future health the level of inequality in 2003 is pro-rich to a similar degree as the realized outcomes in 2002. From 2003 to

<sup>23</sup>We find this is only partly driven by age, i.e. elderly people are more likely to die sooner and tend to be poorer. More discussion is provided in 5.3.5

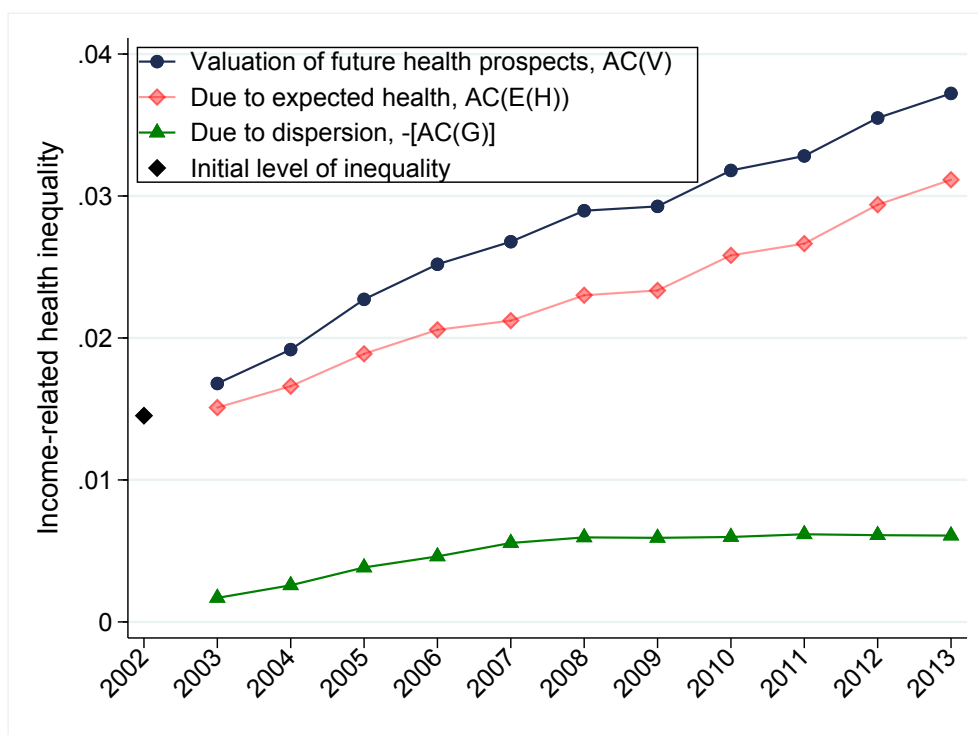


Figure 7: Income-related inequality in the valuation of future health prospects, expected health, and the dispersion

2013 the inequality in expected health rises in a linear fashion such that the inequality in expected health in 2013 is double the current level of cross-sectional inequality observed in 2002. In addition the poor, in general, face greater dispersion in their future health prospects and thus suffer greater losses in the valuation of their future health prospects due to this increased dispersion. The level of inequality in dispersion doubles between 2003 and 2008, but then it stays constant up until 2013. This pattern matches Figure 5b where the dispersion increased at a faster rate for the poor up until 2008, but then increased more equally in absolute terms over the income distribution. Importantly, the additional dispersion among the poor exacerbates the level of inequality in the valuation of future health prospects compared to if one only considered the inequality in expected health.

### 5.3.5. Decomposition of determinants - what factors are associated with the observed inequality in value of future health prospects, expected health and dispersion

To provide a better understanding for the extent to which the variation in each predictor of health is associated with the inequality in the expected health and the dispersion of future health prospects, Figure 8 outlines the contributions from each covariate to the inequality in the valuation of future health prospects and to its two components, expected health and loss in value due to the dispersion of future health prospects. These contributions are based on the absolute concentration indices of all covariates with respect to income rank shown in table 2, the linear regressions of the 1000 predicted quantiles in equation (9) and the decomposition illustrated in equation (18). The main contributors are initial health, initial income, and age.

In general poor people are older and older people have both worse expected health and higher levels of dispersion. The contribution from age also drives the trend of increasing inequality in expected health and partly the trend in the level of dispersion (see Figure D.11 where the contribution of age has been removed by assuming everyone is at the mean age) – although the contribution from age to the dispersion part starts to plateau from 2008 onwards (which also follows the same pattern as the overall dispersion part – the green line/curve in Figure 7). The positive contribution of initial health (which is positively associated with the income rank in 2002) to inequalities in future expected health slightly decreases over time while its positive contribution to the inequality in the dispersion increases

until 2006 before decreasing thereafter. Those who are sicker in the initial period in general also expect to have worse health in the future and face more dispersion in their future health prospects. Even after controlling for these other factors higher initial income still predicts both higher expected future health and lower dispersion in future health prospects. The role of initial income starts from a comparably low position and grows over time, to (jointly with age) overtake initial health as one of the largest contributors to inequality in 2013 future health prospects. The contribution to inequality to expected health is of similar size as initial health in 2013, while the contribution to dispersion is larger from 2006 and onwards.

In comparison, the other factors considered make only very minor contributions to inequality in both expected health and dispersion with the largest contributors being from smoking, where current smokers are poorer, expect to have worse future health and face more dispersion in their future health prospects. An alternative representation of the contribution of the covariates is provided in figure C.9 and C.10 in the appendix, which shows the contribution of each covariate to the inequality in each quantile as illustrated in equation (12).

## 6. Discussion and conclusion

In this paper we have developed a framework to measure income-related inequalities in risky health prospects that considers not only inequalities in expected future health but also inequalities in the dispersion of individual's future health prospects. It complements existing approaches by integrating risk into the standard inequality measurement apparatus used by health economists. It should be used when the social planner wants to account for individual's risk averse preferences in the assessment of income-related health inequalities, provided the social planner has objective information about the distribution of future health outcomes for individuals that have different incomes. We also showed that the approach is equivalent to a weighted average of income-related inequalities for each health quantile, with higher weights given to less favorable quantiles.

We have then illustrated the usefulness of this framework by empirically considering the inequality in future health prospects in an Australian cohort followed up annually from 2002 till 2013. In our empirical study we find that poorer individuals are not only likely to have worse health in the future, on average, but also to face more dispersion, on average, in their future health prospects, which also illustrates the importance of our approach. Age is a major contributing factor with the poor being more likely to be older and older individuals facing both lower expected future health and more dispersion in their future health. However, inequalities are still present also after considering the contribution of age (see Figure D.11). The other main determinants are initial income and initial health. Whereas the contribution of initial income grows the further into the future one looks, the contribution of initial health slightly decreases.

This paper is, as far as we know, the first attempt to integrate risk in health inequality measurement. Our empirical approach is therefore primarily an illustration, and as such there is room for improvements. We used predictions (and not causal estimates) to estimate the conditional distribution of future health prospects and while this improves the accuracy of predicted future health prospects, it also means that the decompositions should not be given a causal interpretation; and rather be seen as conditional associations. A related point concerns omitted predictors in the quantile regressions. Clearly, we do not observe all relevant predictors in our data, and this will increase the estimated risk in future health prospects. Indeed, individuals may have private information that narrow their distribution of future health prospects and this information is more likely to be held by richer rather than poorer individuals. When exploring its impact by also including the explanatory variables observed in 2001, we found only small reductions in the level of risk and it did not change the conclusions of the inequality analysis. In our empirical analyses we also preferred to refrain from statistical inferences as boot-strap procedures would dramatically increase the days required to obtain results.

When deriving this first approach to incorporate risk into health inequality measurement, we have also imposed several assumptions which require scrutiny in future research. First, we assumed our social planner does not care about the impact of any potential mismatches between an individual's subjective risk assessment and the true objective level of risk on income-related inequalities. Such mismatches generally lead to sub-optimal individual decision making and existing literature does suggest that lower educated have less accurate longevity and health expectations (Lange, 2011; Bago d'Uva et al., 2015); therefore only focusing on objective risk – as we do – will most likely provide a lower bound for income-related inequalities. Second, our approach focuses on inequalities in risky health prospects at one future point in time, and thereby neglects health dynamics. While our decomposition goes a long way towards

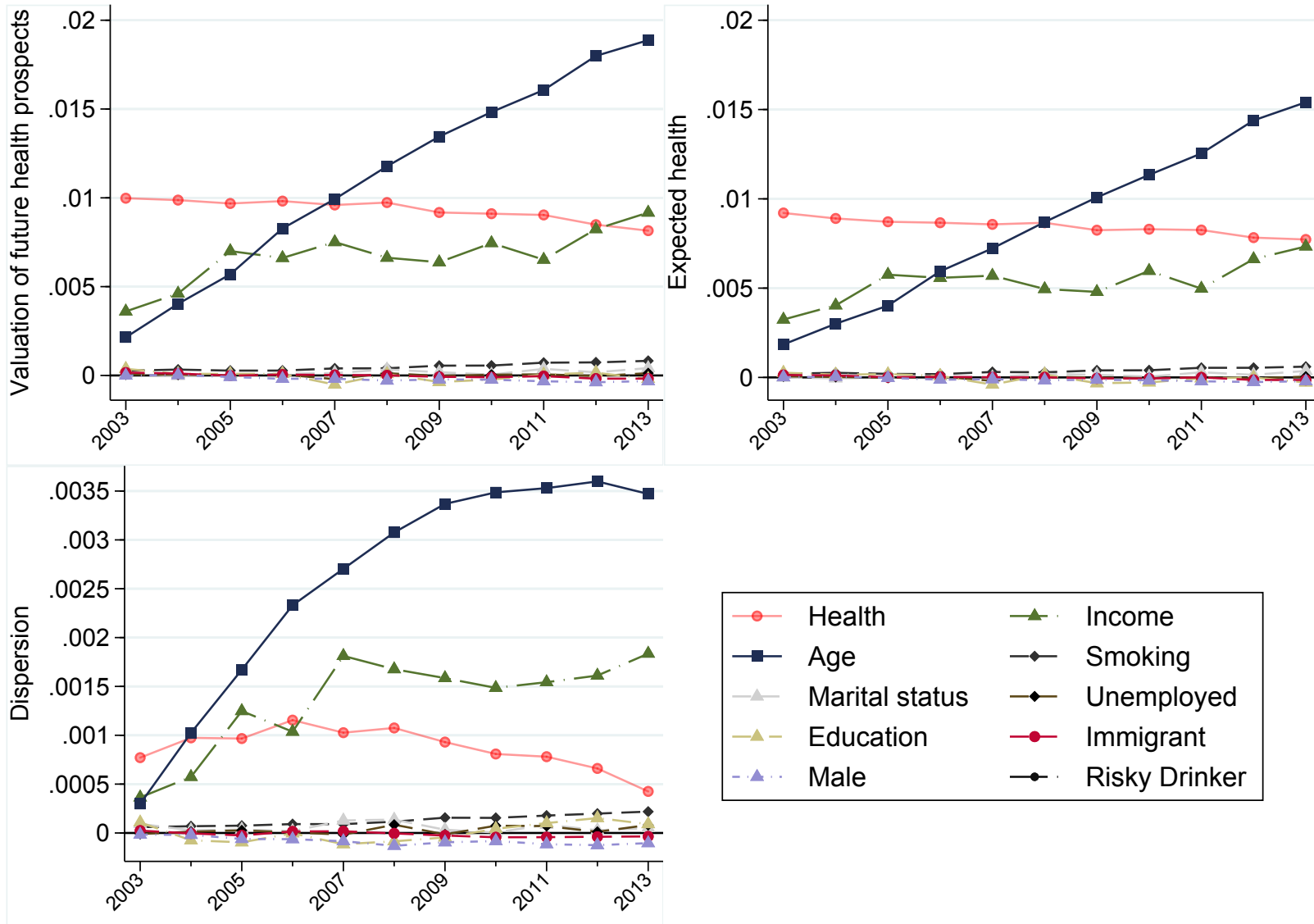


Figure 8: Contributions of key variables to inequalities in the value of the future health prospects, expected health and loss due to dispersion.



such an approach (as it conditions on initial health), a natural future development of our approach is to formally embed the framework in longitudinal decompositions of health inequalities (Van Ourti et al., 2009; Allanson et al., 2010; Allanson and Petrie, 2013b; Baeten et al., 2013), or incorporate the expected stream of future health and the dispersion associated with this stream of health (e.g. lifetime QALYs; duration analyses). Modelling these dynamics and assessing the related inequalities is more complex and also requires a long panel of rich individual data, and although we have such data at hand, these are in general very rare. The fact that our measurement apparatus requires data at only two points in time therefore increases its usefulness considerably.

## Acknowledgment

We would like to thank Aurélien Baillon, Paul Makdissi, Owen O'Donnell, Myra Yazbeck, and participants at conferences and seminars where this paper and earlier versions of this paper were presented, including ASSA in Chicago, Monash University, University of Ottawa, Erasmus University Rotterdam, Vienna University of Economic and Business and the Australian Longitudinal Data Conference 2016. This paper uses unit record data from the Household, Income and Labour Dynamics in Australia (HILDA) Survey. The HILDA Project was initiated and is funded by the Australian Government Department of Social Services (DSS) and is managed by the Melbourne Institute of Applied Economic and Social Research Melbourne Institute. The findings and views reported in this paper, however, are those of the authors and should not be attributed to either DSS or the Melbourne Institute. Petrie was supported under Australian Research Council's Discovery Early Career Awards funding scheme (Project DE150100309). Kjellsson acknowledges Monash University, and the Partner programme at the University of Gothenburg School of Business, Economics, and Law, for supporting a research visit at Monash.

Conflict of interest: The authors have no conflicts of interest to declare.

## References

- Allais, M., 1979. The foundations of a positive theory of choice involving risk and a criticism of the postulates and axioms of the american school (1952), in: Allais, M., Hagen, O. (Eds.), *Expected Utility Hypotheses and the Allais Paradox: Contemporary Discussions of the Decisions under Uncertainty with Allais' Rejoinder*. Springer Netherlands, Dordrecht, pp. 27–145.
- Allanson, P., Gerdtam, U.G., Petrie, D., 2010. Longitudinal analysis of income-related health inequality. *Journal of Health Economics* 29, 78–86.
- Allanson, P., Petrie, D., 2013a. Accounting for population change in the longitudinal analysis of income-related health inequalities, in: Rosa Dias, P., O'Donnell, O. (Eds.), *Health and Inequality*. Emerald Group Publishing Limited. chapter 8, pp. 193–225.
- Allanson, P., Petrie, D., 2013b. Longitudinal methods to investigate the role of health determinants in the dynamics of income-related health inequality. *Journal of Health Economics* 32, 922 – 937.
- Allanson, P., Petrie, D., 2013c. On the choice of health inequality measure for the longitudinal analysis of income-related health inequalities. *Health economics* 22, 353–365.
- Baeten, S., Van Ourti, T., Van Doorslaer, E., 2013. Rising inequalities in income and health in china: Who is left behind? *Journal of health economics* 32, 1214–1229.
- Bago d'Uva, T., Erdogan-Ciftci, E., O'Donnell, O., Van Doorslaer, E., 2015. Who can predict their own demise? accuracy of longevity expectations by education and cognition. *Tinbergen Institute Discussion Paper 52/V*.
- Bleichrodt, H., van Doorslaer, E., 2006. A welfare economics foundation for health inequality measurement. *Journal of Health Economics* 25, 945 – 957.
- Bossert, W., D'ambrosio, C., 2013. Measuring economic insecurity. *International Economic Review* 54, 1017–1030. doi:10.1111/iere.12026.
- Bottai, M., Cai, B., McKeown, R.E., 2010. Logistic quantile regression for bounded outcomes. *Statistics in Medicine* 29, 309–317.
- Brazier, J., Roberts, J., Deverill, M., 2002. The estimation of a preference-based measure of health from the sf-36. *Journal of Health Economics* 21, 271 – 292.
- Case, A., Lubotsky, D., Paxson, C., 2002. Economic status and health in childhood: The origins of the gradient. *American Economic Review* 92, 1308–1334.
- Chou, S.Y., Liu, J.T., Hammitt, J.K., 2003. National health insurance and precautionary saving: evidence from taiwan. *Journal of Public Economics* 87, 1873–1894.
- Cruces, G., Makdissi, P., Wodon, Q., 2004. Poverty measurement under risk aversion using panel data. *B.E. Journal of Economic Analysis & Policy* 3, article 13.
- Currie, J., 2009. Healthy, wealthy, and wise: Socioeconomic status, poor health in childhood, and human capital development. *Journal of Economic Literature* 47, 87–122.
- Cutler, D.M., Lleras-Muney, A., Vogl, T., 2012. Socioeconomic status and health: Dimensions and mechanisms, in: Glied, S., Smith, P.C. (Eds.), *The Oxford Handbook of Health Economics*. Oxford University Press. chapter 7, pp. 124–163.
- Deaton, A., 2013. *The Great Escape. Health, wealth and the origins of inequality*. Princeton university press.
- Drummond, M.F., Sculpher, M.J., Claxton, K., Stoddart, G.L., Torrance, G.W., 2015. *Methods for the economic evaluation of health care programmes*. Oxford university press.

- Duan, N., Manning, W.G., Morris, C.N., Newhouse, J.P., 1983. A comparison of alternative models for the demand for medical care. *Journal of Business & Economic Statistics* 1, 115–126.
- Erreygers, G., 2009. Correcting the concentration index. *Journal of Health Economics* 28, 504–515.
- Erreygers, G., Kessels, R., 2013. Regression-based decompositions of rank-dependent indicators of socioeconomic inequality of health, in: Rosa Dias, P., O'Donnell, O. (Eds.), *Health and Inequality*. Emerald Group Publishing Limited. chapter 14, pp. 227–259.
- Erreygers, G., Van Ourti, T., 2011. Measuring socioeconomic inequality in health, health care and health financing by means of rank-dependent indices: A recipe for good practice. *Journal of Health Economics* 30, 685–694.
- Flores, G., O'Donnell, O., 2016. Catastrophic medical expenditure risk. *Journal of Health Economics* 46, 1–15.
- Gravelle, H., Sutton, M., 2003. Income related inequalities in self assessed health in Britain: 1979–1995. *Journal of Epidemiology & Community Health* 57, 125–129.
- Hagen, O., 1979. Towards a positive theory of preferences under risk, in: Allais, M., Hagen, O. (Eds.), *Expected Utility Hypotheses and the Allais Paradox: Contemporary Discussions of the Decisions under Uncertainty with Allais' Rejoinder*. Springer Netherlands, Dordrecht, pp. 271–302.
- Heckley, G., Gerdtham, U.G., Kjellsson, G., 2016. A general method for decomposing the causes of socioeconomic inequality in health. *Journal of Health Economics* 48, 89–106.
- Jones, A.M., Nicolas, A.L., 2004. Measurement and explanation of socioeconomic inequality in health with longitudinal data. *Health Economics* 13, 1015–1030.
- Kjellsson, G., Gerdtham, U.G., Petrie, D., 2015. Lies, damned lies, and health inequality measurements: Understanding the value judgments. *Epidemiology* 26, 673–680.
- Koenker, R., Bassett, G., 1978. Regression quantiles. *Econometrica* 46, 33–50.
- Kolm, S.C., 1976. Unequal inequalities. ii. *Journal of Economic Theory* 13, 82–111.
- Lange, F., 2011. The role of education in complex health decisions: Evidence from cancer screening. *Journal of Health Economics* 30, 43–54.
- Makdissi, P., Wodon, Q., 2003. Risk-adjusted measures of wage inequality and safety nets. *Economics Bulletin* 9, 1–10.
- Makdissi, P., Wodon, Q., 2018. Poverty-dominant marginal transfer reforms in socially risky situations, in: Greene, W., Khalaf, L., Makdissi, P., Sickles, R., Veall, M., Voia, M. (Eds.), *Productivity and Inequality*. Springer.
- Makdissi, P., Yazbeck, M., 2016. Avoiding blindness to health status in health achievement and health inequality measurement. *Social Science & Medicine* 171, 39–47.
- Mehran, F., 1976. Linear measures of income inequality. *Econometrica* 44, 805–809.
- Melly, B., 2005. Decomposition of differences in distribution using quantile regression. *Labour Economics* 12, 577–590.
- O'Donnell, O., van Doorslaer, E., Wagstaff, E., Lindelow, M., 2008. *Analyzing Health Equity Using Household Survey Data. A guide to techniques and their implementation*. World Bank.
- O'Donnell, O., O'Neill, S., Van Ourti, T., Walsh, B., 2016. conindex: Estimation of concentration indices. *Stata Journal* 16, 112–138.
- Orsini, N., Bottai, M., 2011. Logistic quantile regression in stata. *Stata Journal* 11, 327–344.
- Petrie, D., Allanson, P., Gerdtham, U.G., 2011. Accounting for the dead in the longitudinal analysis of income-related health inequalities. *Journal of Health Economics* 30, 1113–1123.
- Picone, G., Uribe, M., Wilson, R.M., 1998. The effect of uncertainty on the demand for medical care, health capital and wealth. *Journal of health economics* 17, 171–185.
- Portnoy, S., 1991. Asymptotic behavior of the number of regression quantile breakpoints. *SIAM Journal on Scientific and Statistical Computing* 12, 867–883.
- Quiggin, J., 1982. A theory of anticipated utility. *Journal of Economic Behavior & Organization* 3, 323–343.
- Rabin, M., 2000. Risk aversion and expected-utility theory: A calibration theorem. *Econometrica* 68, 1281–1292.
- Rohde, N., Tang, K.K., Rao, D.P., 2014. Distributional characteristics of income insecurity in the u.s., Germany, and Britain. *Review of Income and Wealth* 60, S159–S176. doi:10.1111/roiw.12089.
- Segal, U., Spivak, A., 1990. First order versus second order risk aversion. *Journal of Economic Theory* 51, 111–125.
- Starmer, C., 2000. Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk. *Journal of Economic Literature* 38, 332–382.
- Summerfield, M., Freidin, S., Hahn, M., Li, N., Macalalad, N., Mundy, L., Wooden, M., 2014. *Hilda user manual—release 13* Melbourne: Institute of applied economic and social research. Melbourne: University of Melbourne .
- Tversky, A., Kahneman, D., 1991. Loss aversion in riskless choice: A reference-dependent model. *The quarterly journal of economics* 106, 1039–1061.
- Tversky, A., Kahneman, D., 1992. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty* 5, 297–323.
- Van Doorslaer, E., Van Ourti, T., 2011. Measuring inequality and inequity in health and health care, in: Glied, S., Smith, P.C. (Eds.), *Oxford Handbook of Health Economics*. Oxford University Press, pp. 837–869.
- Van Ourti, T., van Doorslaer, E., Koolman, X., 2009. The effect of income growth and inequality on health inequality: Theory and empirical evidence from the European panel. *Journal of Health Economics* 28, 525–539.
- Van Ourti, T., Erreygers, G., Clarke, P., 2014. Measuring equality and equity in health and health care, in: Culyer, A.J. (Ed.), *Encyclopedia of Health Economics*. Elsevier, San Diego, pp. 234–239.
- Wagstaff, A., van Doorslaer, E., 2000. Chapter 34 equity in health care finance and delivery, Elsevier. volume 1, Part B of *Handbook of Health Economics*, pp. 1803–1862.
- Wagstaff, A., van Doorslaer, E., Watanabe, N., 2003. On decomposing the causes of health sector inequalities with an application to malnutrition inequalities in Vietnam. *Journal of Econometrics* 112, 207–223.
- Wagstaff, A., Paci, P., van Doorslaer, E., 1991. On the measurement of inequalities in health. *Social Science & Medicine* 33, 545–557.
- Wakker, P., 1994. Separating marginal utility and probabilistic risk aversion. *Theory and decision* 36, 1–44.

## AppendixA. Large transfers between different health quantiles of different individuals

Proposition 1 also holds when the health transfer affects the health weights of the unaffected quantiles. Assume that the transfer moves the  $q^r$ th health quantile  $s$  quantiles up and that the  $q^d$ th health quantile of the donating individual moves  $u$  quantiles down. In other words, we are interested in the inequality-impact of individual  $r$  facing  $h_r^1 \leq \dots \leq h_r^{q^r-1} \leq h_r^{q^r+1} \leq \dots \leq h_r^{q^r+s} \leq h_r^{q^r} + b \leq h_r^{q^r+s+1} \leq \dots \leq h_r^Q$  and individual  $d$  facing  $h_d^1 \leq \dots \leq h_d^{q^d-u-1} \leq h_d^{q^d} - b \leq h_d^{q^d-u} \leq \dots \leq h_d^{q^d-1} \leq h_d^{q^d+1} \leq \dots \leq h_d^Q$ . The resulting change of the inequality index equals  $\Delta I = (z_r/nQ) \left\{ \left[ \sum_{l=q^r}^{q^r+s-1} w_l (h_r^{l+1} - h_r^l) \right] + w_{q^r+s} (h_r^{q^r} + b - h_r^{q^r+s}) \right\} + (z_d/nQ) \left\{ w_{q^d-u} (h_d^{q^d} - b - h_d^{q^d-u}) + \left[ \sum_{l=q^d-u+1}^{q^d} w_l (h_d^{l-1} - h_d^l) \right] \right\}$ . This expression is more complex than equation (6), but the same intuition applies. The first term between the  $\{\}$  brackets is always non-negative since  $h_r^{l+1} \geq h_r^l$  and  $h_r^{q^r} + b \geq h_r^{q^r+s}$ ; and the second term between  $\{\}$  brackets is always non-positive as  $h_d^{q^d} - b \leq h_d^{q^d-u}$  and  $h_d^{l-1} \leq h_d^l$ ; just as  $w_{q^r} b$  is always positive and  $-w_{q^d} c$  always negative in equation (6).

**AppendixB. Snapshot of the underlying regressions**

Table B.1: Regression Results

Explanatory variables (2002)	2003						2008						2013								
	Mortality	Health Conditional on survival					Mortality	Health Conditional on survival					Mortality	Health Conditional on survival							
	logistic regression (std. errors)	Logistic quantile coefficients (std. errors)					logistic regression (std. errors)	Logistic quantile coefficients (std. errors)					logistic regression (std. errors)	Logistic quantile coefficients (std. errors)							
	5%	25%	50%	75%	95%	5%	25%	50%	75%	95%	5%	25%	50%	75%	95%	5%	25%	50%	75%	95%	
Health (QALY weight)	0.002 (0.003)	3.430 (0.153)	4.327 (0.086)	4.677 (0.069)	3.918 (0.092)	3.650 (0.153)	0.024 (0.011)	3.247 (0.185)	3.970 (0.102)	4.028 (0.093)	3.200 (0.107)	2.640 (0.196)	0.083 (0.030)	2.459 (0.196)	3.353 (0.109)	3.691 (0.098)	2.840 (0.107)	2.643 (0.197)			
ln(Equiv. Income)	0.913 (0.260)	0.072 (0.029)	0.089 (0.019)	0.064 (0.014)	0.049 (0.017)	0.100 (0.024)	0.754 (0.081)	0.093 (0.031)	0.084 (0.024)	0.068 (0.018)	0.049 (0.019)	0.066 (0.034)	0.878 (0.074)	0.099 (0.039)	0.158 (0.025)	0.122 (0.024)	0.108 (0.022)	0.045 (0.034)			
Age	1.067 (0.074)	-0.002 (0.005)	0.000 (0.003)	-0.002 (0.003)	-0.003 (0.003)	-0.016 (0.006)	1.030 (0.028)	-0.002 (0.006)	0.005 (0.004)	-0.002 (0.003)	-0.009 (0.003)	-0.011 (0.009)	1.009 (0.022)	-0.015 (0.007)	-0.007 (0.005)	-0.005 (0.004)	-0.006 (0.004)	0.006 (0.009)			
Age squared	0.999979 (5.62E-4)	-1.58E-5 (5.37E-5)	-3.50E-5 (3.04E-5)	-2.03E-5 (2.49E-5)	-2.37E-5 (3.56E-5)	5.49E-5 (6.34E-5)	1.0006 (2.25E-4)	-4.40E-5 (6.28E-5)	-1.34E-4 (4.33E-5)	-6.49E-5 (3.58E-5)	5.61E-6 (3.65E-5)	2.68E-6 (9.40E-5)	1.00097 (1.87E-4)	7.93E-5 (7.64E-5)	-4.31E-5 (5.55E-5)	-4.82E-5 (4.52E-5)	-3.19E-5 (4.90E-5)	-1.60E-4 (9.28E-5)			
Male	2.868 (1.066)	0.055 (1.066)	0.044 (0.035)	0.015 (0.021)	0.022 (0.015)	0.055 (0.018)	2.349 (0.036)	0.019 (0.322)	0.049 (0.038)	0.058 (0.025)	0.026 (0.018)	0.073 (0.019)	2.185 (0.042)	0.104 (0.228)	0.123 (0.040)	0.070 (0.030)	0.023 (0.019)	0.074 (0.020)			
Education (Ref: less than year 12)																					
Year12 only	0.562 (0.424)	0.069 (0.424)	0.023 (0.055)	-0.007 (0.031)	-0.026 (0.021)	-0.093 (0.028)	1.170 (0.047)	0.073 (0.254)	0.053 (0.063)	0.047 (0.042)	0.020 (0.029)	0.019 (0.026)	1.062 (0.067)	0.011 (0.183)	-0.028 (0.071)	-0.006 (0.045)	0.013 (0.030)	0.101 (0.031)			
Post School Certificate	1.072 (0.391)	0.059 (0.391)	0.038 (0.046)	0.023 (0.025)	0.010 (0.020)	-0.074 (0.022)	0.832 (0.044)	0.017 (0.128)	0.072 (0.050)	0.039 (0.030)	0.036 (0.023)	-0.008 (0.029)	0.848 (0.052)	0.040 (0.099)	0.035 (0.056)	0.036 (0.037)	0.043 (0.025)	0.053 (0.026)			
Degree or Higher	1.008 (0.534)	0.071 (0.041)	0.031 (0.028)	0.025 (0.019)	0.015 (0.025)	-0.018 (0.048)	1.220 (0.240)	0.154 (0.053)	0.120 (0.031)	0.076 (0.025)	0.062 (0.026)	0.022 (0.059)	0.940 (0.145)	0.147 (0.056)	0.048 (0.037)	0.042 (0.029)	0.053 (0.027)	0.055 (0.059)			
Marital or Civil Status (Ref: Single)																					
Married or living with partner	0.994 (0.650)	0.038 (0.650)	0.054 (0.049)	0.038 (0.032)	0.020 (0.026)	0.004 (0.028)	0.655 (0.056)	0.071 (0.158)	0.038 (0.062)	0.077 (0.037)	0.048 (0.024)	0.004 (0.029)	0.584 (0.064)	0.012 (0.108)	0.057 (0.061)	0.011 (0.041)	-0.021 (0.034)	-0.084 (0.035)			
Divorced/Separated/Widowed	0.961 (0.680)	0.018 (0.680)	-0.002 (0.073)	0.005 (0.044)	0.048 (0.034)	0.194 (0.040)	0.810 (0.069)	0.018 (0.208)	-0.007 (0.078)	0.033 (0.046)	0.019 (0.035)	0.152 (0.047)	0.771 (0.093)	-0.039 (0.153)	0.082 (0.089)	0.003 (0.059)	-0.073 (0.046)	0.011 (0.044)			

Table B.2: Regression Results

Explanatory variables (2002)	2003						2008						2013					
	Mortality	Health Conditional on survival					Mortality	Health Conditional on survival					Mortality	Health Conditional on survival				
	logistic regression (std. errors)	Logistic quantile coefficients (std. errors)					logistic regression (std. errors)	Logistic quantile coefficients (std. errors)					logistic regression (std. errors)	Logistic quantile coefficients (std. errors)				
	5%	25%	50%	75%	95%		5%	25%	50%	75%	95%		5%	25%	50%	75%	95%	
Unemployed	1.932 (2.032)	0.033 (2.032)	-0.014 (0.071)	-0.047 (0.050)	-0.067 (0.047)	-0.115 (0.063)	1.025 (0.117)	-0.117 (0.536)	-0.138 (0.117)	-0.067 (0.085)	0.039 (0.069)	0.210 (0.057)	0.913 (0.152)	-0.284 (0.345)	-0.180 (0.165)	-0.058 (0.099)	0.003 (0.062)	0.104 (0.065)
Smoking status (Ref: Never Smoker)																		
Current Smoker	0.895 (0.457)	-0.178 (0.046)	-0.119 (0.028)	-0.061 (0.021)	-0.051 (0.025)	-0.107 (0.042)	1.202 (0.230)	-0.120 (0.050)	-0.086 (0.031)	-0.079 (0.024)	-0.042 (0.029)	-0.006 (0.051)	2.515 (0.344)	-0.227 (0.059)	-0.100 (0.041)	-0.066 (0.028)	-0.036 (0.027)	-0.015 (0.057)
Former Smoker	1.238 (0.434)	-0.051 (0.040)	-0.048 (0.024)	-0.021 (0.016)	-0.057 (0.021)	-0.069 (0.037)	1.367 (0.183)	-0.010 (0.051)	-0.016 (0.027)	-0.023 (0.019)	0.005 (0.023)	-0.020 (0.047)	1.602 (0.172)	-0.059 (0.046)	-0.056 (0.033)	-0.036 (0.023)	-0.010 (0.023)	0.012 (0.049)
Risky Drinker	0.510 (0.529)	0.007 (0.084)	0.033 (0.039)	0.005 (0.032)	-0.007 (0.037)	-0.030 (0.077)	1.166 (0.349)	0.092 (0.085)	0.030 (0.051)	0.002 (0.031)	-0.046 (0.038)	-0.027 (0.108)	0.845 (0.198)	-0.005 (0.103)	-0.007 (0.058)	-0.015 (0.038)	0.006 (0.048)	0.043 (0.111)
Country of birth (Ref: Australia)																		
Non-English speaking	0.419 (0.257)	-0.144 (0.257)	-0.116 (0.046)	-0.056 (0.035)	-0.050 (0.024)	0.035 (0.030)	0.834 (0.053)	-0.045 (0.154)	-0.008 (0.070)	0.003 (0.041)	-0.028 (0.031)	0.041 (0.028)	0.687 (0.104)	-0.038 (0.101)	0.042 (0.063)	0.016 (0.046)	0.022 (0.031)	0.046 (0.035)
English speaking	1.084 (0.466)	-0.073 (0.060)	-0.006 (0.035)	-0.002 (0.022)	0.034 (0.028)	0.022 (0.057)	0.941 (0.165)	0.040 (0.053)	0.004 (0.036)	0.025 (0.028)	0.013 (0.031)	0.059 (0.059)	0.761 (0.107)	-0.052 (0.080)	0.040 (0.048)	0.050 (0.033)	0.046 (0.030)	0.073 (0.067)
Constant	0.006	-3.171	-3.241	-2.961	-1.875	-0.828	0.025	-3.145	-3.070	-2.498	-1.201	-0.115	0.014	-2.323	-2.652	-2.290	-1.173	-0.474

### Appendix C. Decomposition of each quantile

The graphs in Figures C.9 and C.10 show the contribution of each (set of) covariate to the inequality in each quantile as illustrated in equation (12). Using linear weights, the general decomposition in equation (12) conforms to

$$I(V) = \frac{1}{Q} \sum_{q=1}^Q \frac{(2Q-2q+1)}{Q} \sum_{k=1}^K \beta_k^q AC(x_k) \quad (\text{C.1})$$

The graphs show the (unweighted) contribution of covariate  $k$  to quantile  $q$ . This equals the coefficients  $\beta_k^q$ , which varies over the years, weighted by the absolute concentration index of that covariate  $AC(x_k)$ , which is measured in the initial period 2002. Therefore, these figure also provide the pattern of how the coefficients for each covariate in equation (9) changes over the quantiles.

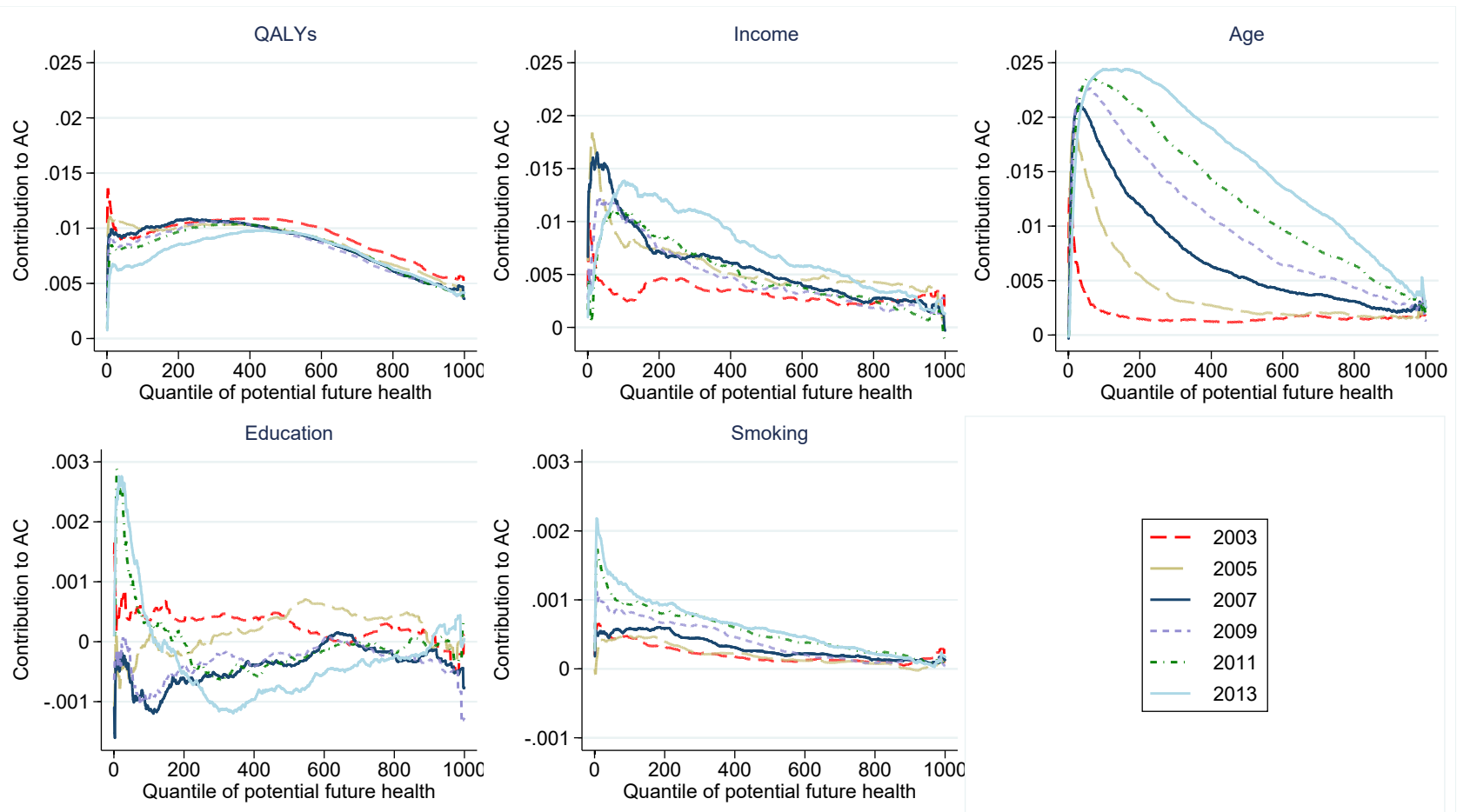


Figure C.9: Contribution to the absolute concentration index per quantile of potential future health: QALYs, Income, Age, Education & Smoking

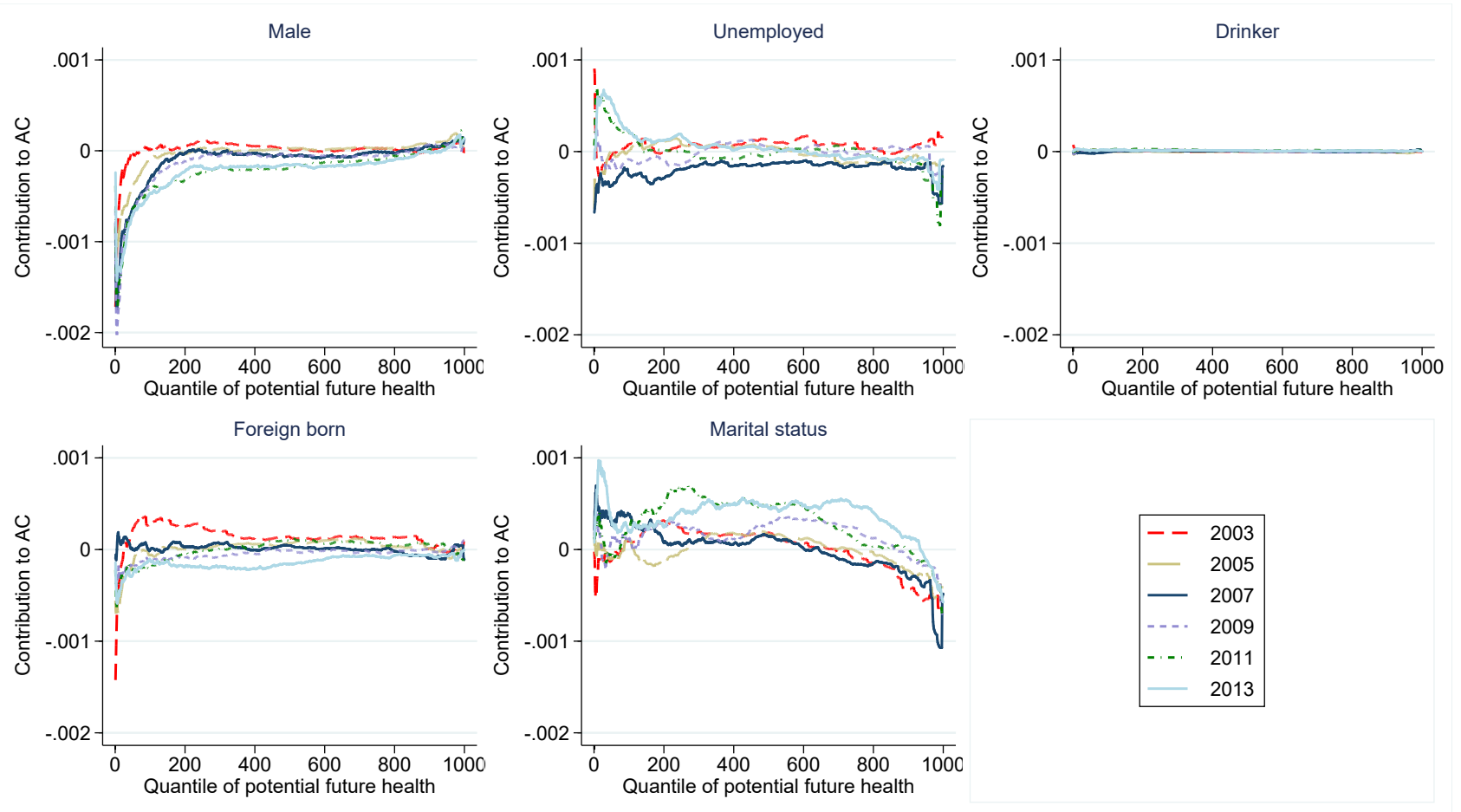


Figure C.10: Contribution to the absolute concentration index per quantile of potential future health: Male, Unemployed, Risky drinker, Foreign Born & Marital Status



**AppendixD. Income-related inequality after controlling for age**

Figure D.11 shows income related inequality after removing the contribution of age using the decomposition in equation (12) (and assuming everyone is at the mean age).

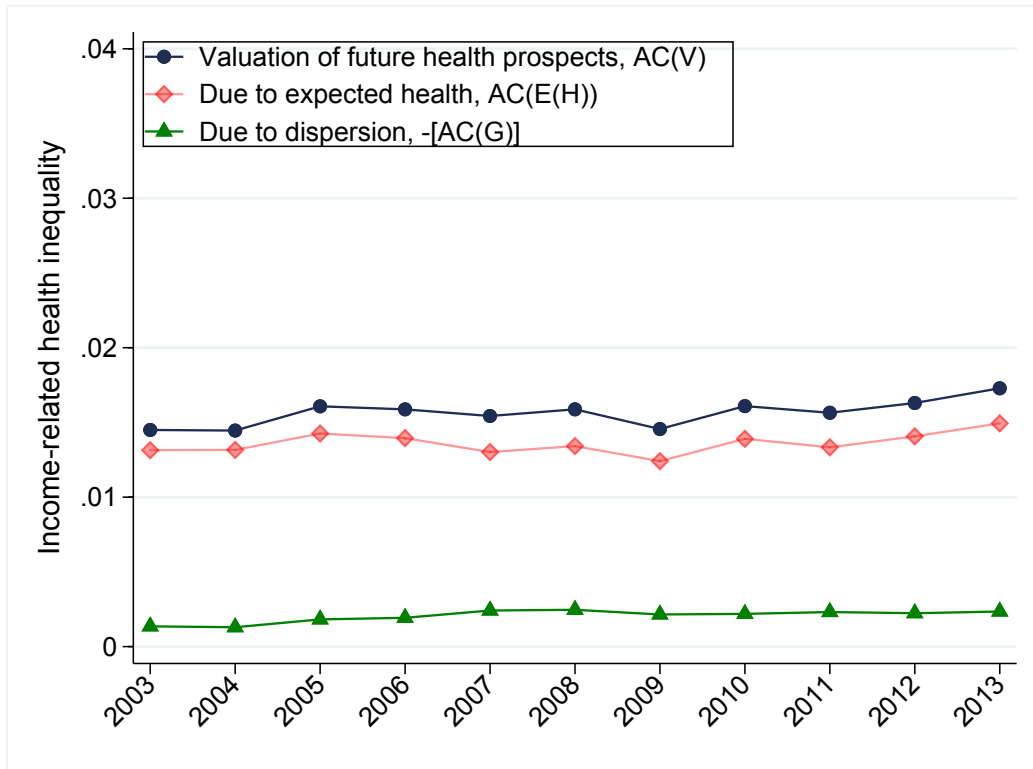


Figure D.11: Income-related health inequality with the contribution of age removed