Published in final edited form as:

Health Econ. 2015 October ; 24(10): 1348–1367. doi:10.1002/hec.3092.

Inequity in the face of death

Pilar García-Gómez^a, Erik Schokkaert^b, Tom Van Ourti^a, and Teresa Bago d'Uva^a

^aErasmus School of Economics, Erasmus University Rotterdam, the Netherlands; Tinbergen Institute, the Netherlands; NETSPAR, the Netherlands ^bCORE, Université Catholique de Louvain, Belgium; Center for Economic Studies, Catholic University of Leuven, Belgium

Abstract

We apply the theory of inequality of opportunity to the measurement of inequity in mortality. Using a rich dataset linking records of mortality and health events to survey data on lifestyles for the Netherlands (1998–2007), we test the sensitivity of estimated inequity to different normative choices and conclude that the location of the responsibility cut is of vital importance. Traditional measures of inequity (such as socioeconomic and regional inequalities) only capture part of more comprehensive notions of unfairness. We show that distinguishing between different routes via which variables might be associated to mortality is essential to the application of different normative positions. Using the fairness gap (direct unfairness), measured inequity according to our implementation of the “control” and “preference” approaches ranges between 0.0229–0.0239 (0.0102–0.0218), while regional and socioeconomic inequalities are smaller than 0.0020 (0.0001). The usual practice of standardizing for age and gender has large effects on measured inequity. Finally, we use our model to measure inequity in simulated counterfactual situations. While it is a big challenge to identify all causal relationships involved in this empirical context, this does not affect our main conclusions regarding the importance of normative choices in the measurement of inequity.

Keywords

equity; equality of opportunities; mortality; lifestyles

1 Introduction

Recently, there has been an increased interest in inequality of opportunity by academics, but the framework has also been receiving growing attention in policy circles such as for example exemplified by the World Development Report 2006 (World Bank, 2006). The theoretical fundamentals of this approach have been explained in important monographs by Roemer (1998) and Fleurbaey (2008). The number of empirical applications is growing rapidly – not only in the domain of income distribution (Roemer *et al.*, 2003; Bourguignon *et al.*, 2007; Devooght, 2008; Lefranc *et al.*, 2008, 2009; Checchi and Peragine, 2010; Aaberge *et al.*, 2011; Almas *et al.*, 2011), but also in other domains such as health (Rosa Dias, 2009; Trannoy *et al.*, 2010; Jusot *et al.*, 2013).

The main intuition of the “inequality of opportunity” approach is that individuals might be held responsible for part of the observed outcome differences. Social concern is then restricted to inequalities in outcomes that are not the responsibility of the individual. Where to draw the line between factors for which individuals should be held responsible and factors for which they should not, is a normative question about which opinions in society differ. Economists have therefore proposed a unifying formal framework, which can accommodate different philosophical perspectives. Suppose the outcome of interest y is linked to a vector of explanatory variables z through the function $y=Y(z)$. The first step is to partition the vector z into variables for which individuals are not held responsible (often called “circumstances” c) and variables for which they are (often called “effort” e). This partitioning reflects a normative choice and can be taken as the starting point for the second step: the derivation of inequality measures which capture the inequality in y in so far (and only in so far) as it is linked to the c -variables. Of course, the results of this second step will depend on the chosen partitioning, but the formal approach can be applied to any partitioning. Note that the traditional focus on pure outcome (e.g. income or health) inequality is not less normative than the other approaches: it is just the special case where considerations of individual responsibility are set aside.

The empirical application of this formal framework raises a difficult issue. While the theoretical rationale for partitioning the vector z is clear, it is often difficult to relate the specific empirical variables in z unambiguously to common opinions about justice or to more sophisticated philosophical theories. Indeed, many variables seem to be of a “mixed” nature, partly reflecting responsibility and choice, partly reflecting circumstances. Lifestyles are a typical example. Roemer (1998, 2002) has put forward an innovative solution to this problem. He proposes to restrict circumstances to a limited set of variables – and he defines individuals to be of the same “type” if they are in identical circumstances, defined in this way. All the other variables are then interpreted as effort. More specifically, he defines the level of effort in terms of the percentile occupied by individuals in the outcome distribution of their “type”. This reflects the normative idea that all variables that are correlated with type should also be interpreted as circumstance. The Roemer-approach gives coherent theoretical foundations for a pragmatic approach to measuring inequality of opportunity. Yet, the narrow definition of “types” gives only a lower bound for a measure of inequality of opportunity. Moreover, the pragmatic stance of Roemer does not take up the challenge of linking the empirical measurement exercise to the rich philosophical debate on the different dimensions of equality of opportunity.

In this paper, we adopt a more ambitious approach. We argue that the difficulty of partitioning z into “effort” and “circumstances” is linked to the fact that variables are related to the observed outcome through different channels, which cannot be disentangled in the reduced form $Y(z)$. In this case, it is essential to estimate a model that distinguishes between these different channels and makes it possible to assess their relative importance (see also Fleurbaey and Schokkaert, 2009).¹ Furthermore, without this richer framework it will not be

¹Deaton (2012) writes: “It is possible that an inequality that might seem to be *prima facie* unjust might actually be the consequence of a deeper mechanism that is in part benevolent, or that is unjust in a different way”. We interpret this quote as an appeal to base equity judgments on the insights of a model that distinguishes between different mechanisms.

possible to implement empirically some particular normative views concerning the location of the responsibility cut (Fleurbaey, 2008). In this paper, we exploit this potential of our approach, and use it to test the sensitivity of inequity measurement to a range of alternative normative choices.

Our empirical application is to mortality. The evaluation of inequalities in the face of death certainly is one of the most ethically challenging issues. In the health domain it is natural to accept that “effort” and individual responsibility are linked to the choice of lifestyles. At least since Grossman (1972), the role of lifestyles has been acknowledged in theoretical models of health production and has also been incorporated in rich models of lifestyle and mortality (e.g. Balia and Jones, 2008, 2011). The identification of effort variables is more difficult in the case of income distribution, since it is hard to simply interpret effort as the number of hours worked. The health setting is therefore an interesting one to test the potential of rich models of inequality of opportunity (Fleurbaey and Schokkaert, 2011). The methodological questions we address - interpretation of a model that distinguishes between different channels, implementation of different philosophical theories – are however relevant for applications of equality of opportunity in all domains.

We exploit rich diagnostic information from hospital admissions and cause of death registries in the Netherlands to estimate a recursive model of whether individuals die during a 7 year time window, considering determinants such as healthy lifestyles and health events. We define our variables so as to avoid reverse causation but there could still be remaining endogeneity due to correlation between unobserved factors affecting lifestyles and health outcomes. Our estimated relationships should therefore be interpreted as associations, rather than as causal effects. However, departure from the necessary assumptions made in our identification strategy will not affect our main conclusions and the main contribution of this paper, namely that the location of the responsibility cut is of vital importance. We will argue in this paper that the estimated associations remain useful in the context of inequity measurement, even if they cannot be interpreted strictly in causal terms.

According to our model, there are strong educational gradients in healthy lifestyles, which in turn are negatively associated with mortality. We use our model to illustrate how different normative views influence the measured degree of inequity. We observe, for example, that the traditional measures of inequity (such as socioeconomic and regional inequalities) only capture part of more comprehensive notions of unfairness. We also show that the usual practice of standardizing for age and gender can have large effects on measured inequity and draw attention to the (often implicit) underlying normative view that inequalities linked to age and gender are fair.

The remainder of the paper is organized as follows. Section 2 introduces our formal framework for measuring inequity. Section 3 describes the data, the econometric model and the estimation results. Section 4 contains our results concerning inequity in mortality, illustrates the importance of the location of the responsibility cut and compares our approach with the Roemer-approach, which does not distinguish between different channels through which mortality is determined. Section 5 concludes.

2 Equality of opportunity and fair allocations

We will introduce our basic measurement concepts for the simplest possible reduced form and then describe how the approach can be extended to a recursive model. The outcome of interest is whether an individual dies during a fixed time window –measured as a binary variable, and termed ‘mortality’ M throughout this text. Mortality is related to a set of variables z through the function $M(z)$. To implement the notion of inequality in opportunity,² we partition the vector z in a subvector of “circumstances” c for which individuals are not held responsible (and that therefore lead to unequal opportunities) and a subvector of “effort” variables e for which individuals are held responsible. Whether an individual dies within the given time window – the mortality of individual i – can then be written as $M_i = M(c_i, e_i)$.

The traditional economic approach, focusing on socioeconomic differences in mortality (see, e.g. Wagstaff *et al.*, 1991; Gerdtham and Johannesson, 2000; van Doorslaer and Gerdtham, 2003; Attanasio and Emmerson, 2003), implicitly takes socioeconomic status (SES) as the circumstance variable –and all other variables as effort. Interpreting lifestyle as the effort variable, and all other variables as circumstances is another option. These are only two possible approaches, and we will describe a whole range of alternative views on the responsibility cut in Section 4. The analysis in this section can be applied to any partition (c_i, e_i) .

For a given partition, the question then is how to measure “illegitimate” inequalities of opportunities, i.e., how to purge a simple inequality measure of the effects for which individuals should be held responsible. Fleurbaey and Schokkaert (2009) propose two methods. The first is called “direct unfairness” – it consists of setting the effort variables at reference values and then measuring inequality in the adjusted advantage measures $a_i = M(c_i, \bar{e})$. It is clear that inequality in a_i can only be due to differences in c_i since the effort variables are fixed. The second starts from the definition of what is considered to be an “equitable situation” in which all inequalities are due to differences in effort. It constructs such an equitable situation by putting all circumstance variables at reference values – and then calculates individual “fairness gaps” (fg_i) as the difference between the actual situation and this equitable point of reference, i.e., $fg_i = M(c_i, e_i) - M(\bar{c}, e_i)$. Inequity is then defined as inequality in these fairness gaps.

The literature has thus far mainly opted (implicitly or explicitly) to calculate direct unfairness. This may indeed seem the most natural of the two approaches. This choice is not innocuous, however. A natural condition to be imposed on an inequity measure is what Fleurbaey and Schokkaert (2009) call “compensation”: inequity should only be zero if two individuals with exactly the same value for the effort (responsibility) variables also have the same mortality. Only in that case do we fully include in our measure of inequity all effects of differences in circumstances. It is immediately clear that the fairness gap satisfies this condition, while the measure of direct unfairness does not. Consider two individuals i and j

²We will use the terms “inequality in opportunity” and “inequity” interchangeably.

with $e_i = e_j = e$. It is very well possible that $M(c_i, \tilde{e}) = M(c_j, \tilde{e})$, while at the same time $M(c_i, e) \neq M(c_j, e)$. On the other hand, $fg_i = fg_j$ if and only if $M(c_i, e) = M(c_j, e)$.

The problem with direct unfairness is linked to the existence of what Gravelle (2003) in a similar setting has called “essential nonlinearities”, i.e. a situation where the value of e influences the marginal effect of c . More generally, the idea of “essential nonlinearities” also helps to understand the differences between the results for direct unfairness and for the fairness gap. Fixing effort (respectively circumstances) at their reference value in $M(c_i, \tilde{e})$ (respectively $M(\tilde{c}, e_j)$) in a certain sense “removes” the impact of the values of e_i (respectively c_j) on the marginal effect of c_i (resp. e_j) on mortality. These “essential nonlinearities” will therefore not be taken into account in the calculation of direct unfairness, which is simply based on $M(c_i, \tilde{e})$. However, they do reappear in the fairness gap through the actual mortality $M(c_i, e_j)$. If the marginal effect of effort on mortality depends positively (negatively) on the value of the circumstances (or vice versa), we may expect a positive (negative) effect on inequity as measured by the fairness gap. For example, and anticipating our results in Sections 3 and 4, we find that mortality during a 7 year time window is higher among older individuals and among individuals with low education, but also that mortality is particularly high among older and lower educated individuals. Under the ethical stance that takes age as the effort variable³ and education as the circumstance variable (commonly known as age-standardisation), the difference between direct unfairness and the fairness gap will crucially depend on the magnitude and sign of this interaction between education and age. Since we find that the combination of low education and old age reinforces the mortality risk, one would expect larger inequities as measured by the fairness gap compared to direct unfairness, which is exactly what we find in our empirical analysis. We will further illustrate the importance of these “essential nonlinearities” in Section 4.⁴

In many cases it will be difficult to classify a given variable z unambiguously in c or in e . The problem can be illustrated –and is at the same time partly solved – by the use of a recursive model that underlies the reduced form $M(z_i)$ and that identifies different channels through which variables are associated with mortality:

$$M_i = m(H_i, L_i, x_i); H_i = h(L_i, x_i); L_i = l(x_i, \pi_i) \quad (1)$$

where M stands for mortality, H for the onset of new health “events (or shocks)”, L for lifestyles, x is a vector of control variables that are related to lifestyles, the onset of health events, and mortality (such as sex or age) and π a vector of preference shifters⁵ (only

³The terminology to call age an “effort” variable may look strange. It just conveys the idea that age differences are seen in this approach as a “legitimate” source of differences in mortality.

⁴The literature (e.g. Fleurbaey, 2008) shows that different approaches to inequality of opportunity not only differ in their degree of respect for the compensation principle, but also embody different ideas about “reward”, i.e. about what is the optimal distribution in the hypothetical situation where all individuals are of the same type. The reward principles underlying direct unfairness and the fairness gap are discussed in Fleurbaey and Schokkaert (2011): they are closely related to the principle of liberal reward.

⁵Preference shifters are individual characteristics that influence mortality or the occurrence of a new health shock only through changes in lifestyles. In the rest of the paper, we consider religion, region and urbanization as preference shifters. Their inclusion is justified in Section 3.2.

influencing L). For expositional purposes it is convenient to summarize eq. (1) in the quasi-reduced form

$$M_i = m[h(l(x_i, \pi_i), x_i), l(x_i, \pi_i), x_i] \quad (2)$$

which can be compared with the reduced form $M_i = M(c_i, e_i)$.

Consider now one specific variable in x_i . Take for the sake of illustration age. Eq. (2) shows that age may be related to mortality through three different channels: (a) age may be related to lifestyles; (b) age may be related to new health shocks (and hence mortality), conditional on lifestyle; and (c) age might be related to mortality, conditional on new health shocks and lifestyle. A priori, there is no reason why these relationships should have the same ethical status, e.g. younger (as compared to older) people could be held responsible for their lifestyle, but maybe less (or not) for the incidence of age on health shocks or mortality. When working with the reduced form these different relationships cannot be distinguished. The finer distinctions can, on the other hand, be introduced in the quasi-reduced form (2). This seems a decisive advantage. We will show relevant examples in Section 4.

Let us finally relax the assumption that the model is fully deterministic. Introducing a stochastic term ε in each of the three equations of eq. (1) yields (in obvious notation) the quasi-reduced form $M_i = M[h(l(x_i, \pi_i, \varepsilon_i^L), x_i, \varepsilon_i^H), l(x_i, \pi_i, \varepsilon_i^L), x_i, \varepsilon_i^M]$. In any empirical application such as our own, one has to estimate the equations in eq. (1). One can then obtain estimates for the mortality risk (or the probability of dying), i.e. the expected value of whether an individual dies during a fixed time window $\hat{E}(M_i/x_i, \pi_i)$ and apply the inequity measures to this. This is the most common approach in applied work. It corresponds to ignoring everything that cannot be explained by the empirical model.⁶

3 Explaining differences in mortality: data and model

We will apply the model in eq. (1) to data for the Netherlands. In this section, we describe the data, the empirical modeling strategy and the resulting estimates.

3.1 Description of the data

We use data from a representative sample of non-institutionalized Dutch individuals taken from the health module of three cross-sectional surveys on living conditions (HSLC hereafter). The HSLC contain information on lifestyles and on the variables x and π . They were linked by Statistics Netherlands to two administrative datasets: the national medical registry (NMR hereafter) and the cause-of-death registry (CD hereafter) which contain, respectively, all hospitalizations between 1998 and 2005, and all deaths between 1998 and 2007 in the Netherlands. These linkages provide us with survival information during 7 years for each individual in the HSLC and with the opportunity to exploit exceptionally rich and

⁶The paper by Lefranc et al. (2009) is an exception in that it distinguishes between observable and unobservable determinants of the variable of interest. The authors also analyze “luck” as a separate variable with its own characteristics, which cannot be reduced to either circumstances or effort.

objective diagnostic information linked to each hospital admission. As far as we know, such information has not been used before to analyze disease-specific impacts of lifestyles on subsequent mortality. Since the linkage between HSLC and NMR is only available from 1998 onwards and since we want to ensure a sufficiently long mortality follow-up, we use the HSLC's for 1998, 1999 and 2000 and ignore more recent waves. We dropped individuals younger than 40 when surveyed by HSLC as they represent only about 5 percent of those who died by 2007.

We model lifestyles, the onset of new health events and subsequent mortality in three stages in chronological order. Indicators of lifestyles, control variables, and preference shifters (respectively, contained in vectors L_t , x_t and π_t) are taken from HSLC ($t = 1998, 1999$ or 2000). New health events H_t are indicators of whether individuals were hospitalized due to the respective health event during the following 5 calendar years (from $t+1$ to $t+5$) but not in t , and are modeled as functions of lifestyles and control variables at time t . Finally, an indicator of whether the individual died between $t+1$ and $t+7$ is assumed to depend on lifestyles and control variables at time t and health events from $t+1$ to $t+5$. The timespan for health events and mortality was determined by data availability (as said above, respectively until 2005 and 2007) and, conditional on this, by our objective to ensure equal time at risk for all individuals (even if for individuals in HSLC 1998 and 1999 a longer follow up was possible). We have a total number of 12,484 observations (individuals). There seems to be sufficient variation in mortality, since 11 percent of the individuals in our sample have died during the considered follow-up period (see Appendix A).

We obtain indicators of the occurrence of *new* health problems from diagnostic information (ICD-9-CM codes) in the NMR, and refrain from using existing health events to minimize biases due to reverse causation. We selected those codes that are likely to correspond to a new health problem if the individual did not go to the hospital for the same code during the previous year. This excludes diagnoses for which individuals are usually first treated as outpatients or which relate to chronic conditions (for example, all mental problems were excluded as these diagnoses are normally related to chronic conditions and are usually preceded by outpatient visits). We considered six groups of new adverse health events: the incidence of i) cancer, ii) circulatory diseases, iii) stroke, iv) respiratory problems, v) digestive problems and vi) genitourinary problems.⁷

We consider three indicators of healthy lifestyles, whether individual: i) is a non-smoker, ii) exercises (at least 1 hour per week) and iii) is not overweight (i.e., if $BMI < 25$). Since we have no information on diet but do control for exercise, the variable “overweight” is intended to proxy for diet. Our vector of control variables x_t is composed of dummy variables representing age-sex categories, highest level of education achieved, home ownership, marital status and whether there are children in the household.⁸ The preference

⁷We are grateful to Isabelle Soerjomataram for her guidance in the identification of new health events from NMR data.

⁸Most papers on (in)equality of opportunity in health focus on the parental background of the individuals. This variable is not available in our data. However, taking parental background as the main “circumstance” variable is usually interpreted as yielding only a lower boundary of the degree of inequality of opportunity. Our model does include many variables (such as education and home ownership) that are associated with parental background and that are interpreted as circumstance variables in some of the normative perspectives we consider.

shifters in vector π_i are indicators of religion, region and urbanization (population density) of the area of residence. More information on all variables used can be found in Appendix A.

3.2 Specification of the model

In this section we explain how we implement empirically the conceptual model in eq. (1). We specify a system of probit equations for whether the individual dies between $t+1$ and $t+7$ (M_i), the new health events H_i between $t+1$ and $t+5$ (cancer, CA ; circulatory diseases, CI ; stroke, S ; respiratory, R ; digestive, D ; and genitourinary disease, G) and indicators L_i at time t of whether the individual does not smoke (NS), exercises (E) and is not overweight (NW).⁹ We specify the following model for the corresponding latent variables $M_i^*, H_{hi}^*, L_{li}^*$

$$M_i^* = \sum_h \beta^h H_{hi} + \sum_l \gamma^l L_{li} + x_i \delta^M + \varepsilon_i^M \quad (3)$$

$$H_{hi}^* = \sum_l \gamma^{hl} L_{li} + x_i \delta^h + \varepsilon_i^h, h = CA, CI, S, R, D, G \quad (4)$$

$$L_{li}^* = x_i \delta^l + \pi_i \lambda^l + \varepsilon_i^l, l = NS, E, NW \quad (5)$$

where x_i and π_i are defined above and observed at time t , and $\beta^h, \gamma^l, \delta^M, \gamma^{hl}, \delta^h, \delta^l, \lambda^l$ are (vectors of) coefficients to be estimated.

As explained above, we observe mortality, new health events and lifestyles at different periods in time. In particular, health events are observed one to five years after lifestyles, and mortality one to seven years after the same variables. By accounting for timing in this way, we avoid reverse causality from the relationship between health events and lifestyles (while obviously this is not an issue in the mortality equation). For example, if we measured exercise and the onset of a respiratory disease in the same year, then there could be reverse causality due to some individuals stopping or starting exercising as a consequence of being diagnosed with a respiratory disease.¹⁰ This may however not be the only source of endogeneity in the model as unobserved factors such as, for example, preferences and frailty, may be correlated with all dependent variables. The resulting potential endogeneity of health events and lifestyles is addressed by assuming that the error terms in our model follow a multivariate normal distribution $MVN(0, \Sigma)$ where Σ is a symmetric matrix with all diagonal

⁹While more detailed information is available concerning survival between t and $t+7$, the exact timing of the new health events, and frequency of smoking, exercise and BMI, we opted to restrict our attention to binary indicators. First, preliminary analysis of our data indicated that these thresholds are the most relevant for mortality and/or health events. Second, this choice makes the estimation of our econometric model more tractable (see also Balia and Jones, 2011).

¹⁰Unfortunately, we only observe lifestyles at t , so we cannot look at effects of average lifestyles that could be a more important determinant of health later in life (Tubef et al., 2012).

elements equal to 1 and off-diagonal elements equal to the correlations between the corresponding error terms.

Identification of correlations between stages of the model requires a valid identification strategy. In our application, this cannot but rely solely on the assumption of correct functional form as our data does not include variables that could be used as exclusion restrictions for the health events equation separately. Our dataset does however contain some variables that might be considered as plausible instruments for the identification of effects of lifestyles on health events *and* mortality equations. Religion has regularly featured as a source of exogenous variation to estimate the impact of lifestyles on economic outcomes (e.g., Auld, 2005),¹¹ while region has been used as an instrument in a model similar to ours (Balía and Jones 2008). We consider therefore that the preference shifters π_i - religion, region and urbanization - only influence lifestyles and not health events (nor mortality). These restrictions are potentially debatable.¹²

A final source of endogeneity might be that we have treated other potentially endogenous variables, such as education and marital status, as exogenous. As soon as one of the identifying assumptions does not hold, estimates of the model in eq. (3)–(5) will not reflect causal estimates. This will limit the usefulness of doing policy or ‘counterfactual’ simulations with our model, but one can argue that the estimated partial associations remain useful for deriving meaningful inequality estimates.

In this respect it is useful to distinguish between the two main causes of endogeneity: reverse causation and omitted variables.¹³ If there is reverse causation between the dependent variable of interest (in our case mortality, health events and lifestyles) and the independent variables, the interpretation of the estimated effect becomes very tricky and basically meaningless. However, in our model we can rather safely exclude the possibility of reverse causation. We do not feel perfectly safe about potential endogeneity through omitted variables, however. Here the consequences for the measurement of inequity will depend on the nature of the variables, and therefore on the normative position taken. Suppose that the omitted underlying variable $\tilde{z}$ is an “illegitimate” one. (The reasoning is analogous when it is “legitimate”). If this variable is correlated with another illegitimate variable c that is included in the model, this invalidates the causal interpretation of the estimated effect of c . However, from the point of view of inequity measurement this is not a problem. On the contrary: it is better that the estimated association between c and the dependent variable of interest reflects to some extent the effect of the equity-relevant but unobservable variable $\tilde{z}$. The situation is less positive when the omitted illegitimate variable $\tilde{z}$ is correlated with a legitimate variable e that is included in the model, as then the overall effect of “legitimate” variables will no longer be measured in an accurate way. However, it seems reasonable to assume that in most cases and for most normative perspectives an omitted illegitimate

¹¹Although, as one referee points out, some studies have shown religion to be a determinant of economic success (Weber, 1905; Barro and McCleary, 2006) which in turn might affect health outcomes.

¹²While not a formal test of instrument validity, univariate probit models for the mortality and health event equations with religion, region and urbanization as additional regressors showed no evidence of joint or individual significance. In addition, the Wald-test of joint significance in the lifestyle equations supports the strength of these instruments ($p < 0.001$).

¹³The same distinction is made by Roemer and Trannoy (2013) in their discussion of the link between causality and equity.

variable will be more strongly associated with the included illegitimate variables than with the included legitimate variables – in fact, while this is not necessary in principle, in practice most normative perspectives will put two variables (whether observable or unobservable) that are strongly correlated with each other in the same category. The measurement of inequity then remains meaningful, even if the effects of the included variables cannot be strictly interpreted in causal terms.

3.3 Estimation results

We estimate the model by maximum simulated likelihood and correct standard errors for clustering at the year and municipality level.¹⁴ The estimated correlations between the equations are shown in Table B.1 in Appendix B. The null hypothesis that all correlation coefficients equal zero (i.e. of a series of univariate probit models) is rejected ($p < 0.0001$). The estimated correlations between health events are mostly positive, reflecting comorbidities¹⁵ (especially between the equations of cancer and other health equations). All correlations between healthy lifestyles are significant. The estimated correlations between mortality/health equations and lifestyles and between mortality and health equations are generally small and insignificant,¹⁶ suggesting that relationships between stages of our model are fully explained by observables. This could be partly due to our definition of health events, which considers only new health events occurring from year $t+1$, while lifestyles are observed in year t , thus avoiding reverse causality.

We now turn to the estimates. In order to assess their magnitude, we present average marginal effects on the probability that each of the outcomes equals 1. For each outcome, we obtain probabilities separately from the respective equations in the multivariate probit model. This corresponds to neglecting the correlations between error terms of mortality and each of the health events (lifestyles) and health events and lifestyles, which were generally small and insignificant. Additionally, we obtain total marginal effects of lifestyles on the probability of dying within 7 years. For each healthy lifestyle, these total effects are obtained by: i) computing, for each individual, the marginal probabilities of suffering each of the health events, with and without adopting the lifestyle, ii) replacing the actual occurrence of each health event by these probabilities, i.e. by their expected values; and iii) computing the resulting marginal probabilities of dying with and without adopting the lifestyle.

Table 1 presents average marginal effects for each equation in columns 3 to 12 and total effects of lifestyles on the probability of dying in column 2. With respect to the mortality equation, we find a positive and significant impact of the health events upon mortality, with the exception of genitourinary and circulatory problems.¹⁷ The highest marginal effect on

¹⁴Maximum simulated likelihood is consistent when the number of observations is large and when there is a sufficient number of draws. In practice, a reasonably small number has been found to result in well-behaved estimates. The estimates presented in this paper are based on five draws. Five draws lead to an optimization time of one week and increasing the number further would drastically increase optimization time of our 10-equation model (Hajivassiliou, 1997). Therefore, we tested the sensitivity of our estimates to the number of draws by experimenting with sub-models of our 10-equation model which take considerably less time to optimize. We monitored the behavior of the estimates when increasing the number of draws from 5 to 250 and found that our main results remained qualitatively unchanged.

¹⁵12.4% of those with at least one new health event experience two or more new health events.

¹⁶As a consequence, the equations of the model could have been estimated separately (at the cost of some loss of efficiency due to the larger and significant correlations mentioned above). One is however only able to conclude this after estimation of the full model that estimates correlations freely.

mortality is that of cancer which increases the probability of dying within 7 years by about 14 percentage points, while digestive problems, stroke and respiratory problems increase that probability between 7 and 10 percentage points. Conditional on the observed health events and control variables, we still find that healthy lifestyles are negatively related to mortality. Not smoking or exercising decreases the probability of dying by about 3 percentage points, conditional on the health problems considered here, while the further decrease caused by not being overweight is small and insignificant. The control variables have a significant, but smaller marginal effect, with the exception that elderly (and especially elderly males) have a much higher probability of dying (for example, the probability of dying within 7 years for males older than 80 is 57 percentage points higher than that of females aged 50–60).

The estimated marginal effects for the health event equations show that lifestyles are related to mortality through prevention of adverse health events, namely stroke (not smoking and exercising significantly decrease the probability of having a stroke within 5 years, see column “Stroke”) and digestive problems (not being overweight and not smoking significantly decrease the probability of having digestive problems within 5 years, see column “Digest”). Thus, being overweight is related to mortality only via the onset of health events, while, as we have seen above, the other lifestyles also have a direct impact on mortality. We find no association between lifestyles on the other health events which was unexpected for cancer (all lifestyles are insignificant in the column “Cancer”).¹⁸ Contrary to the mortality equation, we find that the control variables are rather unimportant for the onset of new health events. We find a few significant but small marginal effects. The only exception is age and gender which play a more prominent role, capturing disease-specific age-gender patterns (age-gender dummies are highly significant in health equations, see columns “Cancer” to “Genito”, except for digestive problems). The remaining variables have hardly any direct relationship with the occurrence of health events, over and above the indirect effect that they may have via lifestyles.

Finally, the control variables are more important in the lifestyle equations, showing for example a strong educational gradient in healthy lifestyles. Preference shifters (religion, region and urbanization) also contribute significantly to the explanation of lifestyle differences.

4 Inequity in mortality risks: the importance of the responsibility cut

In this section we focus on inequity in mortality *risks* (or the probability of dying between $t+1$ and $t+7$). We will (i) show how different (c, e)-partitionings –reflecting different normative views – influence the measurement of inequity; (ii) discuss the results of a set of policy-relevant counterfactual simulations and iii) compare our results with those of the Roemer-approach, which does not disentangle the different channels through which variables might be related to mortality.

¹⁷The finding for circulatory diseases is in line with the observation that most deaths from cardiovascular diseases occur before patients arrive at the hospital.

¹⁸We tried to check whether the absence of any effect remains when restricting to lifestyle-related cancers, but the number of observations with lifestyle-related cancers is too limited.

Using the complete recursive model, we simulate for each individual his/her predicted probability of dying, conditional on the actual values of all variables (more details in Section 3.3). Call this M_i^B . To measure “direct unfairness”, we additionally simulate probabilities $M^S(c_i, \tilde{e})$ conditional on actual values of circumstance variables and reference values of effort variables. Finally, the fairness gap is computed as $fg_i = M_i^B - M^S(\tilde{c}, e_i)$ where $M^S(\tilde{c}, e_i)$ is the simulated probability of dying conditional on actual values of effort variables and reference values of circumstances. All these calculations neglect the estimated correlations between the error terms of the multivariate probit and the actual mortality experience of individuals, i.e., whether M_i equals 0 or 1.

In order to calculate the different measures, we have to choose reference values $\tilde{c}$ and $\tilde{e}$. For the fairness gap it is reasonable to pick as reference values for $\tilde{c}$ the characteristics of the type that can be assumed to be in the “best” situation. This fits in the interpretation of $M(\tilde{c}, e_i)$ as an equitable reference point.¹⁹ For consistency, we then make a similar choice for $\tilde{e}$. The reference values corresponding to the “best” situation are obtained as the categories for each of the control variables x_i and the preference shifters π_i that have the lowest probability of dying, conditional on the remaining variables, as predicted from our multivariate probit model.²⁰

4.1 An overview of different normative choices

To structure our discussion, we write the quasi-reduced form (2) explicitly in terms of the variables that have been introduced in the previous section, $M_i = m[h(l(ed_i, ho_i, d_i, ag_i, r_i, b_i), ed_i, ho_i, d_i, ag_i) l(ed_i, ho_i, d_i, ag_i, r_i, b_i), ed_i, ho_i, d_i, ag_i]$ where ed_i stands for education, ho_i for home ownership, ag_i for age and gender, d_i for being married and having children, r_i includes region and urbanization and b_i stands for religious beliefs. In the notation used earlier, the preference shifters are $\pi_i = (r_i, b_i)$. Different normative perspectives can now easily be accommodated within this framework. Table 2 shows different partitions of the set of variables into legitimate (effort) and illegitimate (circumstance) sources of inequality. We explain in detail the rationale behind all these partitions in the rest of this section.

The left panel of Table 3 shows the corresponding results for inequity with the fairness gap, and the right part those with direct unfairness.²¹ As a measure of inequality we use the variance.²² The evaluation of the actual “baseline” situation is in columns 2 (using fairness gap) and 7 (using direct unfairness). The remaining columns refer to the results of

¹⁹It is also in line with the broader analysis in Fleurbaey and Schokkaert (2011) where the choice of $\tilde{c}$ is related to the position taken with respect to vertical equity in health care.

²⁰We use as reference values: i) those living in a less densely populated area, ii) in the West of the Netherlands, iii) married individuals, iv) with children; v) home owners; vi) those who have completed at least an undergraduate (non-) university degree, and vii) Presbyterians. Only for the age-gender categories, have we deviated from this procedure by choosing the youngest females as the reference category, rather than those between 50 and 60 which have a slightly lower estimated probability of dying.

²¹Bootstrapped confidence intervals for differences in inequity estimates confirm the statistical relevance of our findings. Results are available from the authors.

²²Our outcome, predicted probability of dying, is bounded by 0 and 1. This boundedness has consequences for our choice of inequality measure. Erreygers (2009) has proposed the idea that inequality in attainments (here: predicted probability of dying) should be equal to inequality in shortfalls (here: predicted probability of surviving). This so-called “complementarity”-property was further explored by Lambert and Zheng (2011), who show that the variance is the only measure satisfying this requirement that is also subgroup decomposable.

counterfactual simulations and will be discussed in the next subsection. The last row in Table 3 gives the mean predicted probability of dying – this is the average of M_i^B .

A useful benchmark is that of “pure” inequality in mortality risks M_i^B , which can be interpreted as the case where all differences are considered to be illegitimate. In that case direct unfairness and the fairness gap coincide – see the first row in Table 3.²³ As soon as we accept that individuals are held responsible for some variables, inequity (or inequality of opportunity) does no longer coincide with pure inequality. How to think about this individual responsibility? Two broad streams can be distinguished in the literature on responsibility-sensitive egalitarianism (Fleurbaey, 2008; and Fleurbaey and Schokkaert, 2011, for a discussion in the context of health).

The original philosophical inspiration of that literature (Rawls, 1971; Dworkin, 1981) was that persons as autonomous moral agents should assume responsibility for their goals and their conception of the good life, i.e., that they should be held responsible for their preferences. This “preference approach” was criticized by authors such as Arneson (1989), Cohen (1989), and Roemer (1998). They claimed that preferences are often the product of social influences, for which individuals cannot be held responsible, and they advocated the common-sense view that individuals should be held responsible only for what they have genuinely chosen, as opposed to what they have inherited from circumstances. This “control” (or choice) approach has dominated the empirical literature until now, perhaps because it indeed captures common sense. Yet, it is not so easy to implement as it may seem. Indeed, from a broader ethical perspective, “genuine control” requires correcting for interindividual differences in the environment and also for differences in the decision-making abilities of the individuals. But this brings us on a slippery slope: if, as scientists, we reason within a deterministic model, what room is left for control? Where do we have to stop in our quest for underlying causes that are not under the control of the individual? Does free choice really exist?²⁴ Note that this philosophical discussion is linked to the questions about causality that were raised in the previous section.

Although these considerations may look very abstract, they have to be faced when operationalizing the control-approach. Indeed, for each of the variables in the quasi-reduced form, one has to decide if it is the result of individual choice or not. Age and gender certainly are not under individual control, but what about educational attainment? This is partly a matter of choice, partly a matter of innate (uncontrolled) capacities. For the purposes of this exercise we assume that educational attainment is not a matter of choice. All the other variables are considered to be under the control of the individuals.²⁵ This means that the advantage measure for direct unfairness in the control approach becomes

$$a_i^{CONTROL} = m[h(l(ed_i, \tilde{h}o, \tilde{d}, ag_i, \tilde{r}, \tilde{b}), ed_i, \tilde{h}o, \tilde{d}, ag_i), l(ed_i, \tilde{h}o, \tilde{d}, ag_i, \tilde{r}, \tilde{b}), ed_i, \tilde{h}o, \tilde{d}, ag_i], \text{ while the}$$

²³This is due to defining the fairness gap as an absolute deviation and measuring its inequality by the variance, a translation-invariant inequality measure.

²⁴For economists working within the paradigm of rational choice, the problem of determining the boundaries of control is even more difficult. Indeed, in this paradigm individual decisions result from a mechanical optimization exercise with a given objective (preferences) and a given set of options (determined by the budget set and possibly additional constraints). All endogenous variables are causally influenced by other variables of the model and therefore cannot embody free will (Fleurbaey 2008, p. 87).

²⁵Classifying religion, region and urbanization as circumstance hardly changes the results.

fairness gaps are given by

$$fg_i^{CONTROL} = M_i^B - m[h(\widetilde{ed}, ho_i, d_i, \widetilde{ag}, r_i, b_i), \widetilde{ed}, ho_i, d_i, \widetilde{ag}], l(\widetilde{ed}, ho_i, d_i, \widetilde{ag}, r_i, b_i), \widetilde{ed}, ho_i, d_i, \widetilde{ag}].$$

The “preference approach” (PREF) holds individuals responsible for their preferences, i.e., their conceptions of a good life, even if these preferences are not chosen and are not under their control. At first sight, this is easier to implement in our setting, as we can simply consider that individuals are responsible for their lifestyle (but not for the additional factors influencing health shocks or mortality). The expressions for a_i^{PREF} and fg_i^{PREF} (and all other ethical positions in this section) can be easily derived using the same logic as above.

Holding individuals fully responsible for their lifestyle is perhaps too simplistic, however. First, chosen lifestyles reflect both preferences and environmental factors (e.g. the budget constraint). A theoretically coherent implementation of the “preference”-approach would therefore be to assume constrained utility maximization, specify a functional form for lifestyle preferences and identify its parameters through the estimation of the structural model. While our data did not allow this more ambitious approach,²⁶ we may still be able to correct to some extent for economic constraints on lifestyle choices. Educational attainment is a good proxy for these constraints. Second, the philosophical argument for holding individuals responsible for their preferences is that these reflect their *authentic* views of the good life. Involuntary addictions and biased information should in this richer view be corrected for. Again, particularly interesting questions arise with respect to the effect of education – it has been argued by Cutler and Lleras-Muney (2010) that, correcting for income, the effect of education on lifestyles is mainly related to differences in cognitive abilities. Could we then not draw the conclusion that these do not reflect differences in genuine preferences? Here again, we arrive on a slippery slope: when does the correction of revealed preferences become unacceptable paternalism? Yet, to illustrate the implications of both concerns (the economic and the philosophical one), we define a third ethical position (“authentic preferences”, PREFA), where the effect of education on lifestyles is considered illegitimate. Note that we need our recursive model to implement the (“authentic”) preference approach, as we have to distinguish between the roles of some variables in the different equations of the model.

The results for these different approaches are given in the “baseline” columns of Table 3. In the case of direct unfairness, the differences are substantial. Note the much smaller value for control, and the substantial effect of purging preferences of education (compare “PREFA” and “PREF”). The differences are much smaller for the fairness gap. Indeed, holding individuals responsible for variables under control or for lifestyles hardly decreases this inequity measure, compared to the case where all variables are illegitimate sources of inequality. The distinction between direct unfairness and the fairness gap turns out to be vital: “essential nonlinearities” are crucially important. We explained already in section 2 why we should indeed expect a larger value for inequity based on the fairness gap in the case

²⁶Moreover, specifying a functional form for preferences would necessarily require restricting the range of possible heterogeneity in individual preferences.

where the marginal effects of the circumstance/effort variables depend positively on the value of the effort/circumstance variables.

The “control” and “preference” approaches have strong philosophical underpinnings. The economic literature, however, has until now focused on more pragmatic and partial approaches such as socioeconomic or regional inequalities (Lee, 1991; Smith, 1999; Wagstaff and van Doorslaer, 2000; Cutler *et al.*, 2006; Bengtsson and van Poppel, 2011). Moreover, standardization for age and gender is quite common. This follows from the idea that differences are only inequitable if they are caused by institutions – and that inequalities linked to age and gender reflect irremediable biological differences (Wagstaff and van Doorslaer, 2000).

Pure inequality *after standardization for age and gender* can almost be seen as a second benchmark of comparison. The results in Table 3 (row “STAND”) show that this standardization has a tremendous effect on measured inequity, which falls considerably – more so for direct unfairness than for the fairness gap, where it remains more substantial (0.0061). Note, however, that the normative status of this standardization exercise is far from clear. Age and gender are not under the control of the individuals – and they may influence preferences. The reference to “irremedial” inequality versus that caused by “institutions” is not very convincing either: surely the relation between age and gender on the one hand, and health on the other hand is not invariant over time and space and does depend on policy.

The partial approaches that are prominent in the health economic literature accept that standardization is desirable and focus on the inequality due to socioeconomic status and region. It is easy to translate them in our framework. For the former we take $c=(ed, ho)$,²⁷ for the latter we take $c=r$ (i.e., region and urbanization). The resulting measures of inequity turn out to be very small. The differences with the control-approach, which gives a much larger inequity-value, are particularly striking. This is not so surprising for region, as this is only significant in the lifestyle equations. It is less straightforward for socioeconomic status, however, as education plays an important role in the explanatory model. We come back to the role of education in the next subsection.

The results in Table 3 show that the decision to classify age and gender either as legitimate or illegitimate sources of inequality is crucial. This makes sense since age and gender are the most important determinants of mortality (consult the average marginal effects in Table 1). To investigate the importance of this effect, we re-evaluated the “preference” and “control” approaches with age as a legitimate source.²⁸ We find that inequities become very small for direct unfairness. The inequities decline but remain much larger than those for socioeconomic inequality for the fairness gap which is in line with our earlier explanation of “essential nonlinearities” implying here that the marginal effect of education is stronger for older individuals. It is important to emphasize that these findings should *not* be interpreted as a weakness of our approach. Quite the contrary, there are good reasons for the differential

²⁷We treat home ownership as an indicator of wealth and therefore of socio-economic status. A sensitivity analysis with socio-economic status only measured by education yields very similar results.

²⁸These results are available from the authors on request.

treatment of age and gender in the different normative approaches, and our findings simply point to the crucial importance of the choices made in this regard.

4.2 Simulating counterfactual situations

Additional insights into the interplay between the different underlying variables and different normative perspectives can be obtained from counterfactual simulations. We consider, in that order, the consequences of changing the distribution of education, of changing lifestyles and of removing gender differences. Results are shown in columns 3 to 6 and 8 to 11 in Table 3. Our discussion will focus on equity, but useful additional interpretations can be derived from the mean predicted probability of dying M_i^B in the counterfactual situations, as given in the last row of Table 3.

When some of our identifying assumptions are violated (see section 3.2 for additional discussion), we can no longer interpret the results presented in this section as counterfactual simulations that might help policy makers in addressing inequality of opportunity in mortality. Strictly speaking, such policy evaluation requires estimates of causal effects of all variables that are considered exogenous in this model but that are potentially endogenous to health outcomes and lifestyles, such as education (see e.g., Van Kippersluis et al., 2011). However, our results remain useful as an illustration of how to measure and compare inequality of opportunity in two different states of the world and of the importance of the normative choices.

4.2.1 Education and inequity—Educational differences are one of the driving forces of inequity in health and changing the distribution of education is often seen as an important component of any attempt to improve the health situation of the population (see, e.g., Conti et al., 2010; Johnson, 2010; Oreopoulos and Salvanes, 2011).²⁹ We therefore simulate two counterfactual situations. The first (educ1) consists in equalizing education for all at the highest level. This is not a realistic goal, but the results give us an idea on the upper bound of equity and efficiency that can be reached by changing the distribution of education. The second simulation (educ2) is perhaps more realistic. It raises the educational attainment of the lowest educated group to the second lowest education level. As mentioned before, these simulations should be interpreted as illustrations only: while our recursive model addresses potential endogeneity in the relationship between health events and lifestyles, it assumes that education is an exogenous variable. This is sufficient to illustrate the usefulness of our approach to “inequality of opportunity” in section 4.1, but may be too restrictive for simulating counterfactual educational situations.

Not surprisingly, simulation educ2 has a smaller effect on the mean predicted probability of dying than educ1. There is a substantial change in the measured inequity in both simulations, both for direct unfairness and for the fairness gap—and this *whatever the ethical stance* that is taken. (Of course, one has to take into account that regional and socioeconomic inequity was already small in the actual situation and therefore cannot improve much in absolute terms). This result sheds light on the results concerning socioeconomic inequality

²⁹There has also been interest in the effects of education policy on wage inequality (De Fraja, 2001; Brunello, Fort and Weber, 2009).

(where SES is operationalized by education and home ownership) in the previous subsection. It would be very misleading to conclude from the latter results that education is after all not so important from the point of view of equality of opportunity. Quite the contrary, the counterfactual simulations show that education is an important driving force of inequity in the face of death. The reason again has to be found in the “essential nonlinearities”: the education effect differs with the level of other important variables, such as age and gender.

It is interesting to compare the results for PREF and for PREFA. Remember that in the partitioning PREF, lifestyle differences are considered to lead to legitimate inequalities, even if they are explained by educational differences. With the partitioning PREFA this is not the case. One would expect that equalizing education has a much stronger influence on inequity in the latter case. This is exactly what is found with direct unfairness. It is not true for the fairness gap, however. As mentioned before, this is because the fairness gap includes all indirect effects of circumstances in the measure of inequity.

4.2.2 Changing lifestyles—Suppose now that society succeeds in having all individuals exercise at least one hour a week.³⁰ The results are given in columns 5 and 10 (“Exercise”). There is a surprisingly large decline in average probabilities of dying, which is even larger than in the ambitious educ1-simulation that assigned to all individuals the highest possible educational level.

The equity effects of this counterfactual simulation are interesting, because responsibility for lifestyles was one of the main factors differentiating the “preference” from the “control” approach. This has strong effects in the case of direct unfairness. If we hold people fully responsible for their level of exercise (approach PREF) inequity is *not* affected by this counterfactual simulation – for obvious reasons. In the base situation direct unfairness is calculated as the variance in $M^S(c_i, \tilde{e})$ – where $\tilde{e} = e^*$, the “best” possible lifestyle. The counterfactual “exercise” situation simulates the predicted probabilities for $e_i = e^*$. In the control approach, however, equalizing the level of exercise has a large effect on measured inequity, as we assume that many of the factors influencing the level of exercise are beyond individual control. Note, however, that there is also a strong effect in the case of PREFA: in this approach we do not hold individuals responsible for differences in exercise that are associated with differences in educational level. This leads to a larger perceived inequity in the base situation, but this difference between PREF and PREFA becomes smaller if exercise is equalized.

Changing lifestyles has also strong effects in the fairness gap approach, but here they are similar for the various ethical positions –for the reasons explained earlier, this is most surprising for PREF. Again, the explanation lies in the importance of the “essential nonlinearities”.

4.2.3 Removing gender differences—Let us finally consider a simulation in which we remove the gender gap, i.e., we equalize the lifestyles, the occurrence of health shocks and

³⁰We also simulated (i) the eradication of smoking, (ii) the eradication of overweight, (iii) lifestyle prevalence of the late 1980s, but do not show the results because they had similar (and somewhat smaller) efficiency and equity impacts. All results are obtainable from the authors upon request.

the direct effect on mortality for men and women (columns 6 and 11 – “Gender”). This is not a realistic short run goal, but there are indications in the literature that the gender gap is recently becoming smaller –and that gender differences are (at least to some extent) influenced by social factors (Rogers *et al.*, 2010; Quah, 2011). Surely the different treatment of men and women in society will have an impact on the gender gap. The counterfactual where the gender gap disappears may be illustrative for the potential importance of this effect, or, at least, gives an idea about its upper bound.

Of all counterfactual simulations, removing gender differences is associated with the strongest decline in the average probability of dying. It is also associated with a strong decrease of inequity (measured both with the fairness gap and with direct unfairness) for all the ethical approaches that do not hold individuals “responsible” for their gender – i.e., for the “philosophically inspired” preference and control-approaches. As soon as one accepts the need for standardization, however, the effect on inequity is much smaller (see the results for STAND). Since we know that gender differences are not fully biologically determined, there seems to be a real issue here. The common practice of quasi-automatic demographic standardization may yield a biased picture of inequity.

4.3 The Roemer-approach

The previous sections have illustrated the advantages of working with the full recursive model. It makes it possible to differentiate the channels through which variables impact mortality. At the same time it requires a lot of information which is not always available. It is therefore interesting to compare our approach with the more popular, pragmatic, approach proposed by Roemer (1998). As described before, the latter approach consists of defining “types” as individuals with the same values of the circumstance variables and then comparing the outcomes of these types. Effort variables are deliberately *not* included, since individuals are considered to exert the same effort when they lie at the same rank in the distribution of mortality for their circumstance type. All variables which are correlated with explicitly defined circumstances are also implicitly considered as circumstances, as they are not freely chosen by the individuals.

We compare the Roemer method with our control-approach in which $c_i = (ag_i, ed_i)$. In its most basic version (see e.g. Rosa Dias, 2009), the Roemer-method is empirically implemented by regressing mortality upon the type characteristics, i.e., estimating $M_i = \tilde{m}(ag_i, ed_i)$, without including any other explanatory variables. This is a deliberately misspecified model. The “misspecification” is justified as a way to take up the effect of effort variables that are correlated with circumstances in the estimated coefficients of the circumstance variables. One then computes direct unfairness as inequality in the simulated values of $\tilde{m}(ag_i, ed_i)$, again omitting the stochastic part.

We have implemented this approach upon our data by estimating a single “reduced” probit equation for mortality with age-gender and education as the only explanatory variables. The misspecification of the model changes the coefficient estimates drastically (compare the average marginal effects in Table 1 and Appendix C). This leads to a direct unfairness estimate of 0.0228, much larger than the estimate of 0.0102 for direct unfairness within the control approach based on our recursive model. Given the results described before, it is not

surprising that integrating the correlation between legitimate and illegitimate variables into the measure of inequity leads to a higher value for the inequity measure. In fact, the Roemer-approach gives a value which is very close to our results with the fairness gap. In this sense, it seems to be a useful approximation in the case where effort variables are not available and the ethical stance does not distinguish between different channels driving mortality.

The Roemer-approach raises some normative issues, however. By construction it picks up in the circumstances all correlations with effort variables. This makes it impossible to accommodate normative positions where this correlation does not necessarily lead to illegitimate inequity. The most prominent example is the preference approach: obviously preferences can be correlated with circumstances (e.g. with socioeconomic status), but this would not mean that individuals should not be held responsible for them (or, formulated differently, one can still argue that preferences should be respected, whatever their origin). If one wants to implement such normative positions, in which some sources of correlation matter for defining inequality of opportunity while others do not, the pragmatic solution of Roemer will not be sufficient and one cannot do without the estimation of a more refined model.

5 Conclusion

In this paper we focused on inequity in mortality. We estimated – on a rich dataset – a model that identifies different channels through which variables are related to death during a 7 year time window. We have used this model to implement different approaches and measures from the theory of inequality of opportunity. Rather than just summarizing the findings of our empirical work, we draw attention to some methodological and conceptual issues that point to useful directions for future research and that go beyond the specific issue of inequity in mortality.

(a) The traditional measures of inequity that are most popular among economists (such as socioeconomic or regional inequalities) should not be interpreted as measuring a comprehensive notion of unfairness. They only capture a part of inequality in opportunity. The recent theories of equality of opportunity have introduced a formal framework which can be used to accommodate richer normative views. This makes it possible to link the empirical literature more closely to the cut between legitimate and illegitimate sources of inequality that has been suggested by different philosophical theories. The flexible nature of the recent economic approach of equality of opportunity allows for meaningful sensitivity analyses to compare the implications of these different normative perspectives.

(b) The usual practice of standardizing for age and gender in health economic applications should be reconsidered. It has a tremendous effect on measured inequity. In so far as demographic (mainly gender) differences are codetermined by social and behavioral factors and are not only linked to biological differences, they should be considered explicitly in any analysis of inequality of opportunity.

(c) The pragmatic approach proposed by Roemer (1998) is very useful in many cases, but is not sufficiently flexible to integrate relevant approaches such as, e.g., the preference

approach to inequality of opportunity. If the available data are sufficiently rich, it is worthwhile to go beyond the estimation of a reduced form. This is needed to differentiate between different normative approaches, since the same reduced form variable may work through different channels that have different normative implications. Moreover, nonlinearities in the model may be essential. The difference between direct unfairness and the fairness gap is therefore of crucial importance.

(d) To get an idea about the relative importance of the different channels through which individual and environmental characteristics have an impact on the dependent (equity-relevant) variable, it is not necessary that all the effects can be interpreted strictly in causal terms. Associations with a common underlying unobservable variable can also be indicative of inequity. Yet, our recursive model is only a first step in the direction of a full structural model, and it should be the long run ambition to collect better data for estimating richer structural models. This is definitely necessary when one accepts the preference approach to inequality of opportunity, as this requires the identification of the preference parameters from a full-fledged model of utility maximization.

(e) Counterfactual simulations are useful to get a better insight into the relative importance of different explanatory factors. Moreover, they show how the evaluation of the equity of counterfactual situations depends on the normative position that is taken. However, the results of these simulations should only be used for policy advice if the assumptions made to identify causal relationships in the econometric model are valid.

Acknowledgments

Teresa Bago d'Uva, Pilar García-Gómez, and Tom Van Ourti are supported by the NETSPAR project "Health and income, work and care across the life cycle II". Pilar García-Gómez also acknowledges support from a Marie Curie fellowship, and Tom Van Ourti from the National Institute on Ageing, under grant R01AG037398. Teresa Bago d'Uva and Pilar García-Gómez are Postdoctoral Fellows of the Netherlands Organisation for Scientific Research–Innovational Research Incentives Scheme–Veni. We thank the Netherlands Central Bureau of Statistics for access to the linked data sets used for this research (POLS Gezondheid 1998–2000, LMR 1998–2005, DO 1999–2007), and Isabelle Soerjomataram for valuable help with identifying new health events from NMR data.

References

- Aaberge R, Mogstad M, Peragine V. Measuring long-term inequality of opportunity. *Journal of Public Economics*. 2011; 95:193–204.
- Almas I, Cappelen A, Lind J, Sorensen E, Tungodden B. Measuring unfair (in)equality. *Journal of Public Economics*. 2011; 95(7–8):488–499.
- Arneson R. Equality and equal opportunity for welfare. *Philosophical Studies*. 1989; 56:77–93.
- Attanasio O, Emmerson C. Mortality, health status and wealth. *Journal of the European Economic Association*. 2003; 1(4):821–850.
- Auld M. Smoking, drinking, and income. *Journal of Human Resources*. 2005; 15:504–518.
- Balia S, Jones A. Mortality, lifestyle and socio-economic status. *Journal of Health Economics*. 2008; 27:1–26. [PubMed: 17448554]
- Balia S, Jones A. Catching the habit: a study of inequality of opportunity in smoking-related mortality. *Journal of the Royal Statistical Society A*. 2011; 174(1):175–194.
- Barro R, McCleary R. Religion and Economy. *Journal of Economic Perspectives*. 2006; 20:49–72.
- Bengtsson T, van Poppel F. Socioeconomic inequalities in death from past to present: an introduction. *Explorations in Economic History*. 2011; 48:343–356.

- Bourguignon F, Ferreira F, Menéndez M. Inequality of opportunity in Brazil. *Review of Income and Wealth*. 2007; 53(4):565–618.
- Brunello G, Fort M, Weber G. Changes in compulsory schooling, education and the distribution of wages in Europe. *The Economic Journal*. 2009; 119:516–539.
- Checchi D, Peragine V. Inequality of opportunity in Italy. *Journal of Economic Inequality*. 2010; 8:429–450.
- Cohen G. On the currency of egalitarian justice. *Ethics*. 1989; 99:906–944.
- Conti G, Heckman J, Urzua S. The education-health gradient. *American Economic Review: Papers & Proceedings*. 2010; 100:234–238.
- Cutler D, Deaton A, Lleras-Muney A. The determinants of mortality. *Journal of Economic Perspectives*. 2006; 20(3):97–120.
- Cutler D, Lleras-Muney A. Understanding differences in health behaviors by education. *Journal of Health Economics*. 2010; 29(1):1–28. [PubMed: 19963292]
- Deaton, A. What does the empirical evidence tell us about the injustice of health inequalities?. In: Nordheim, Ole Frithjof, et al., editors. *Inequalities in health: ethics and measurement*. Oxford University Press; 2012.
- De Fraja G. Education policies: equity, efficiency and voting equilibrium. *Economic Journal*. 2001; 111:C104–C119.
- Devooght K. To each the same and to each his own. A proposal to measure responsibility sensitive income inequality. *Economica*. 2008; 75(298):280–295.
- Dworkin R. What is equality? Part 2: Equality of resources. *Philosophy and Public Affairs*. 1981; 10:283–345.
- Erreygers G. Can a single indicator measure both attainment and shortfall inequality? *Journal of Health Economics*. 2009; 28(4):885–893. [PubMed: 19409631]
- Fleurbaey, M. Fairness, responsibility and welfare. Oxford: Oxford University Press; 2008.
- Fleurbaey M, Schokkaert E. Unfair inequalities in health and health care. *Journal of Health Economics*. 2009; 28(1):73–90. [PubMed: 18829124]
- Fleurbaey, M., Schokkaert, E. Inequity in health and health care. In: Barros, P.McGuire, T., Pauly, M., editors. *Handbook of Health Economics*. Vol. 2. New York: Elsevier; 2011. p. 1003-1092.
- Gerdtham U-G, Johannesson M. Income-related inequality in life-years and quality-adjusted life-years. *Journal of Health Economics*. 2000; 19(6):1007–1026. [PubMed: 11186842]
- Gravelle H. Measuring income related inequality in health: standardisation and the partial concentration index. *Health Economics*. 2003; 12:803–819. [PubMed: 14508866]
- Grossman M. On the concept of health capital and the demand for health. *Journal of Political Economy*. 1972; 80:223–255.
- Hajivassiliou, V. STICERD: Econometrics Paper Series from Suntory and Toyota International Centres for Economics and Related Disciplines. 1997. Some practical issues in maximum simulated likelihood.
- Johnson RC. The health returns of education policies from preschool to high school and beyond. *American Economic Review: Papers & Proceedings*. 2010; 100:188–194.
- Jusot F, Tubeuf S, Trannoy A. Circumstances and Efforts: How important is their correlation for the measurement of inequality of opportunity in health? *Health Economics*. 2013 (in press).
- Lambert P, Zheng B. On the consistent measurement of achievement and shortfall inequality. *Journal of Health Economics*. 2011; 30(1):214–219. [PubMed: 21129796]
- Lee C. Regional inequalities in infant mortality in Britain, 1861–1971: patterns and hypotheses. *Population Studies*. 1991; 45(1):55–65. [PubMed: 11622608]
- Lefranc A, Pistolesi N, Trannoy A. Inequality of opportunities vs inequality of outcomes: are Western societies all alike? *Review of Income and Wealth*. 2008; 54(4):513–546.
- Lefranc A, Pistolesi N, Trannoy A. Equality of opportunity and luck: definitions and testable conditions, with an application to income in France. *Journal of Public Economics*. 2009; 93:1189–1207.
- Oreopoulos P, Salvanes KG. Priceless: The nonpecuniary benefits of schooling. *Journal of Economic Perspectives*. 2011; 25(1):159–184.

- Quah S. Gender and the burden of disease in ten Asian countries: an exploratory analysis. *Asia Europe Journal*. 2011; 8:499–512.
- Rawls, J. A theory of justice. Cambridge, Mass.: Harvard University Press; 1971.
- Roemer, J. Equality of opportunity. Cambridge, Mass.: Harvard University Press; 1998.
- Roemer J. Equality of opportunity: a progress report. *Social Choice and Welfare*. 2002; 19(2):455–471.
- Roemer J, Aaberge R, Colombino U, Fritzell J, Jenkins S, Lefranc A, Marx I, Page M, Pommer E, Ruiz-Castillo J, San Segundo M, Tranaes T, Trannoy A, Wagner G, Zubiri I. To what extent do fiscal regimes equalize opportunities for income acquisition among citizens? *Journal of Public Economics*. 2003; 87:539–565.
- Roemer, J., Trannoy, A. Equality of opportunity. Cowles Foundation Discussion Paper No. 1921. In: Atkinson, A., Bourguignon, F., editors. *Forthcoming in Handbook of Income Distribution*. Vol. 2. Amsterdam: Elsevier (North-Holland); 2013.
- Rogers R, Everett B, Saint Onge J, Krueger P. Social, behavioral, and biological factors, and sex differences in mortality. *Demography*. 2010; 47(3):555–578. [PubMed: 20879677]
- Rosa Dias P. Inequality of opportunity in health: evidence from a UK cohort study. *Health Economics*. 2009; 18:1057–1074. [PubMed: 19644964]
- Smith J. Healthy bodies and thick wallets. *Journal of Economic Perspectives*. 1999; 13(2):145–166.
- Trannoy A, Tubeuf S, Jusot F, Devaux M. Inequality of opportunities in health in France: a first pass. *Health Economics*. 2010; 19(8):921–938. [PubMed: 19588460]
- Tubeuf S, Jusot F, Bricard D. Mediating role of education and lifestyles in the relationship between early-life conditions and health: Evidence from the 1958 British Cohort. *Health Economics*. 2012; 21(SUPPL. 1):129–150.
- van Doorslaer E, Gerdtham U-G. Does inequality in self-assessed health predict inequality in survival by income? Evidence from Swedish data. *Social Science and Medicine*. 2003; 57(9):1621–1629. [PubMed: 12948571]
- van Kippersluis H, O'Donnell O, Van Doorslaer E. Long run returns to education: Does schooling lead to an extended old age? *Journal of Human Resources*. 2011; 46:695–721.
- Wagstaff A, Paci P, van Doorslaer E. On the measurement of inequalities in health. *Social Science and Medicine*. 1991; 33(5):545–557. [PubMed: 1962226]
- Wagstaff, A., van Doorslaer, E. Equity in health care finance and delivery. In: Culyer, A., Newhouse, J., editors. *Handbook of Health Economics*. Vol. 1B. Amsterdam: Elsevier (North-Holland); 2000. p. 1803-1862.
- Weber, M. *The Protestant ethic and the spirit of capitalism*. London & Boston: Unwin Hyman; 1905. (published in 1930)
- World Bank. *World Development Report 2006: Equity and Development*. Washington, D.C.: The World Bank; 2006.

Appendix

A - Description of the data

Table A.1

Variable description

Variable	Mean
Dependent variables	
Died between t+1 and t+7	0.110
Health problems diagnosed between t+1 and t+5	
Cancer	0.055
Circulatory disease	0.044

Variable	Mean
Stroke	0.023
Respiratory disease	0.013
Digestive disease	0.011
Genitourinary disease	0.010
Healthy lifestyles	
Not overweight (BMI<25)	0.483
Nonsmoker currently	0.705
Exercise (more than 1 hour a week)	0.456
Control variables	
Age-gender	
Male between 40 and 50	0.160
Female between 40 and 50 (reference category)	0.179
Male between 50 and 60	0.137
Female between 50 and 60	0.142
Male between 60 and 70	0.098
Female between 60 and 70	0.099
Male between 70 and 80	0.063
Female between 70 and 80	0.072
Male 80+	0.019
Female 80+	0.030
Married	0.751
Has children	0.251
Highest degree of education attained	
Primary education (reference category)	0.256
Lower vocational education	0.196
Lower general or scientific secondary education	0.090
Higher vocational, general or scientific secondary education	0.271
Higher vocational education or a university degree	0.188
Home owner	0.631
Preference shifters	
Religion	
Catholic	0.359
Protestant - Dutch reformed	0.174
Protestant - Presbyterian	0.074
Other protestant (reference category)	0.327
Muslim	0.009
Other religion, or not religious	0.057

Variable	Mean
Region	
North	0.120
East	0.231
West (reference category)	0.404
South	0.245
Urbanization	
Very low population density (reference category)	0.178
Low population density	0.240
Average population density	0.221
High population density	0.242
Very high population density	0.120

The 6 groups of adverse health events were obtained as follows. First, we aggregated all diagnoses that could indicate new adverse health events in the following groups: i) infectious (infectious and parasitic diseases –some codes within 001–139); ii) cancer (neo-plasms – some codes within 140–239); iii) endocrine (endocrine, nutritional and metabolic diseases and immunity disorders –some codes within 240–279); iv) nervous (diseases of the nervous system and sense organs –some codes within 320–389); v) circulatory (some codes within 390–422 within diseases of the circulatory system); vi) stroke (some codes within 430–459 within diseases of the circulatory system); vii) respiratory (diseases of the respiratory system–some codes within 460–519); viii) digestive (diseases of the digestive system –some codes within 520–579); ix) genitourinary (diseases of the genitourinary sys-tem –some codes within 580–629); x) skin (diseases of the skin and subcutaneous tissue –some codes within 680–709); xi) musculoskeletal (diseases of the musculoskeletal system and connective tissue –some codes within 710–739); xii) injury (injury and poisoning –some codes within 800–999). Second, we estimated a univariate mortality model and excluded those groups showing no evidence of influence, i.e., groups iii), iv), x) and xii). Third, we removed all groups with an incidence below 1%, i.e., groups i) and xi) to avoid too small cell sizes. More information can be obtained from the authors upon request.

B - Estimation Results

Table B.1

Estimated correlation coefficients between equation of multivariate probit model

	Dead	Cancer	Circulat.	Stroke	Respir.	Digest.	Genito.	Nonsmo.	Exerc.
Cancer	0.035								
Circulatory	0.077	0.025							
Stroke	−0.014	0.082 *	0.050						
Respiratory	0.076	0.164 ***	0.052	0.101 *					
Digestive	−0.033	0.091 *	0.053	0.011	−0.119 *				
Genitourinary	0.068	−0.044	0.128 *	0.009	0.078	0.118 *			

	Dead	Cancer	Circulat.	Stroke	Respir.	Digest.	Genito.	Nonsmo.	Exerc.
Nonsmoker	0.012	0.032	−0.006	0.003	−0.018	0.034	0.011		
Exercise	0.000	0.009	0.010	−0.035	0.004	0.029	−0.028	0.151***	
Not overweight	0.044	0.020	−0.007	0.014	0.022	0.046	−0.016	−0.137***	0.069***

***, ** and * represent significance at 1%, 5% and 10%, respectively.

C - Some results Roemer model

Table C.1

Marginal effects in the "reduced" Roemer model

Inequity	0.023
Marginal effects	
Male between 40 and 50	0.000
Male between 50 and 60	0.039
Female between 50 and 60	0.006
Male between 60 and 70	0.130
Female between 60 and 70	0.047
Male between 70 and 80	0.346
Female between 70 and 80	0.181
Male 80+	0.499
Female 80+	0.653
Lower vocational	−0.025
Low general	−0.035
Higher secondary	−0.033
Higher education	−0.048

Table 1

Estimated marginal effects derived from the multivariate probit model.

Total effects of lifestyles		Effects in individual equations ^c										
	Dead	Dead	Cancer	Circulat.	Stroke	Respir.	Digest.	Genito.	Nonsmoker	Exercise	Not overweight	
Health events												
Cancer		0.137***										
Circulatory		0.010										
Stroke		0.094***										
Respiratory		0.098*										
Digestive		0.074*										
Genitourinary		0.025										
Lifestyles												
Nonsmoker	-0.036 ^b	-0.031***	-0.007	-0.009	-0.016***	-0.003	-0.009***	0.003				
Exercise	-0.037 ^b	-0.034***	-0.008	-0.008	-0.006*	-0.004	-0.003	-0.002				
Not overweight	-0.009 ^b	-0.006	-0.008	-0.007	-0.005	-0.003	-0.008***	0.001				
Age/gender ^d												
Male 40-50		0.000	-0.032***	0.010**	0.002***	-0.001	-0.002	-0.007***	-0.045***	-0.031**	-0.155***	
Male 50-60		0.031***	-0.012*	0.033***	0.009	0.003	0.000	-0.006***	-0.003	-0.103***	-0.219***	
Female 50-60		0.001	-0.006	0.007*	0.002	-0.002	-0.004	0.010***	0.095***	-0.007	-0.070***	
Male 60-70		0.100***	0.039***	0.080***	0.032***	0.015***	0.000	0.003	0.100***	-0.100***	-0.163***	
Female 60-70		0.036***	0.007	0.028***	0.011***	0.003	-0.001	0.019***	0.197***	-0.027	-0.135***	
Male 70-80		0.273***	0.072***	0.107***	0.064***	0.025***	0.002	0.010**	0.132***	-0.157***	-0.077***	
Female 70-80		0.138***	0.024***	0.052***	0.037***	0.015***	0.006	0.020***	0.293***	-0.110***	-0.104***	
Male 80+		0.571***	0.016	0.044***	0.021***	0.050***	0.014	-0.002	0.231***	-0.299***	-0.051	
Female 80+		0.400***	0.012	0.041***	0.064***	0.026***	0.008	0.006	0.358***	-0.261***	-0.071**	

Effects in individual equations ^c										
Total effects of lifestyles										
	Dead	Dead	Cancer	Circulat.	Stroke	Respir.	Digest.	Genito.	Nonsmoker	Exercise
Education^a										
Lower voc.	-0.015**	-0.002	-0.002	-0.002	-0.002	0.003	0.000	-0.001	0.048***	0.077***
Lower gen.	-0.026**	-0.001	0.004	-0.001	-0.002	0.003	0.002	0.004	0.057***	0.125***
Higher sec.	-0.017**	-0.006	0.002	-0.006	-0.001	-0.003	0.002	0.008***	0.072***	0.162***
Higher educ.	-0.021**	-0.014**	-0.002	-0.014**	-0.004	0.003	-0.001	-0.002	0.132***	0.275***
Home owner	-0.019***	-0.004	-0.005	-0.004	-0.001	-0.003	-0.002	-0.003*	0.075***	0.096***
Married	-0.024***	0.001	0.002	0.001	0.001	-0.002	-0.001	-0.001	0.078***	0.043***
Has children	-0.029***	-0.012**	-0.012**	-0.001	-0.006	-0.004	-0.004	0.012***	0.018	-0.037***
Religion^a										
Catholic									0.014	0.022*
Dutch reformed									0.034***	-0.043***
Presbyterian									0.069***	-0.033*
Muslim									0.115***	-0.190***
Other/none									0.069***	-0.062***
Region^a										
North									-0.033**	-0.036**
East									-0.025**	-0.012
South									-0.032***	-0.015
Urbanization^a										
Very high									-0.047***	-0.00428
										0.039**

Total effects of lifestyles		Effects in individual equations ^c										
		Dead	Dead	Cancer	Circulat.	Stroke	Respir.	Digest.	Genito.	Nonsmoker	Exercise	Not overweight
High										-0.027***	-0.004	0.012
Average										-0.023*	0.003	0.027*
Low										-0.013	0.009	0.020
Mc Fadden's R ²		0.004										

^aReference categories described in Appendix A.

^bWe have not calculated the statistical significance of total effects

^c***, ** and * represent significance of the corresponding estimated coefficient in the multivariate probit model at 1%, 5% and 10%, respectively

Table 2

Partitioning between legitimate and illegitimate sources of inequality (effort and circumstances) for different ethical positions

	Legitimate	Illegitimate
All Illegitimate (ALLILLEG)		Age-gender, education, married, children, home ownership, religion, region and urbanization
Control (CONTROL)	Married, children, religion, region, urbanization, home ownership	Age-gender, education
Preference (PREF)	Lifestyles	Age-gender, education, married, children, home ownership, for health events and mortality
Authentic preference (PREFA)	Age-gender, home ownership, married, children, religion, region, urbanization (i.e., all but education) for lifestyles	Education for lifestyles, health events and mortality; age-gender, married, children, home ownership, for health events and mortality
Standardization (STAND)	Age-gender	Education, married, children, religion, region, urbanization, home ownership
SES Inequality (SES)	Age-gender, married, children, religion, region, urbanization	Education, home ownership
Regional inequality (REG)	Age-gender, education, married, children, religion, home ownership	Region, urbanization

Inequity in mortality according to different ethical positions using the fairness gap and direct unfairness

Table 3

Ethical position	Fairness gap					Direct unfairness				
	Baseline	Counterfactual simulations			Baseline	Educ1	Educ2	Exercise	Gender	Gender
		Educ1	Educ2	Exercise						
<i>ALLILEG</i>	0.0239	0.0183	0.0215	0.0167	0.0158	0.0183	0.0215	0.0167	0.0158	0.0158
<i>CONTROL</i>	0.0229	0.0174	0.0206	0.0161	0.0149	0.0072	0.0089	0.0064	0.0058	0.0058
<i>PREF</i>	0.0239	0.0182	0.0215	0.0168	0.0157	0.0121	0.0131	0.0146	0.0094	0.0094
<i>PREFFA</i>	0.0239	0.0182	0.0215	0.0167	0.0157	0.0158	0.0192	0.0170	0.0148	0.0148
<i>STAND</i>	0.0061	0.0031	0.0047	0.0041	0.0047	0.0001	0.0001	0.0001	0.0002	0.0002
<i>SES</i>	0.0020	0.0005	0.0012	0.0011	0.0015	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
<i>REG</i>	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
$\overline{M}_i^B$	0.1039	0.0850	0.0971	0.0814	0.0761	0.0850	0.0971	0.0814	0.0761	0.0761