Mind the Gap: A Study of Cause-Specific Mortality by Socioeconomic Circumstances

Daniel H. Alai* Séverine Arnold (-Gaille)†
Madhavi Bajekal‡ Andrés M. Villegas§

Abstract

Socioeconomic groups may be exposed to varying levels of mortality; this is certainly the case in the United Kingdom, where the gaps in life expectancy, differentiated by socioeconomic circumstances, are widening. The reasons for such diverging trends are yet unclear, but a study of cause-specific mortality may provide rich insight into this phenomenon. Therefore, we investigate the relationship between socioeconomic circumstances and cause-specific mortality using a unique dataset obtained from the U.K. Office for National Statistics. We apply a multinomial logistic framework; the reason is twofold. First, covariates such as socioeconomic circumstances are readily incorporated, and, second, the framework is able to handle the intrinsic dependence amongst the competing causes. As a consequence of the dataset and modeling framework, we are able to investigate the impact of improvements in cause-specific mortality by socioeconomic circumstances. We assess the impact using (residual) life expectancy, a measure of aggregate mortality. Of main interest are the gaps in life expectancy among socioeconomic groups, the trends in these gaps over time, and the ability to identify the causes most influential in reducing these gaps. This analysis is performed through the investigation of different scenarios: first, by eliminating one cause of death at a time; second, by meeting a target set by the World Health Organization (WHO), called WHO 25×25; and third, by developing an optimal strategy to increase life expectancy and reduce inequalities.

Keywords: Cause-of-Death Mortality Data, Socioeconomic Circumstances, Multinomial Logistic Regression, Cause Elimination, Life Expectancy

*Centre for Actuarial Science, Risk and Investment, School of Mathematics, Statistics and Actuarial Science, University of Kent, Canterbury, Kent CT2 7NF, United Kingdom, d.h.alai@kent.ac.uk.

†Department of Actuarial Science, Faculty of Business and Economics, University of Lausanne, 1015 Lausanne, Switzerland, severine.arnold@unil.ch

‡Department of Applied Health Research, Institute of Epidemiology and Health Care, University College London, London, WC1E 6BT, United Kingdom, m.bajekal@ucl.ac.uk

§School of Risk and Actuarial Studies and ARC Centre of Excellence in Population Ageing Research (CEPAR), UNSW Business School, UNSW Sydney, NSW 2052, Australia, a.villegas@unsw.edu.au

1 Introduction

Over the past 30 years, life expectancy in high-income countries has increased dramatically, averaging a gain of about two years per decade. However, these gains have not been shared equally across all groups in society. In England, people living in the poorest residential areas can, on average, expect to die eight years earlier than people living in the most affluent (Office for National Statistics (2015a)). Recent evidence indicates that inequalities in life expectancy in England have not only widened (Office for National Statistics (2015b)), but are forecast to widen even further (Bennett *et al.* (2015); Villegas and Haberman (2014)). This situation is not unique to England; elsewhere in Europe, notwithstanding differences in the size and trend of the absolute gap in (standardized) mortality rates between the most and least advantaged social groups across countries, the underlying message is the same: health inequalities are ubiquitous and have persisted over time (Mackenbach *et al.* (2016); Brønnum-Hansen and Baadsgaard (2012)).

In the United Kingdom, and elsewhere in Europe, the goals of public health policy were redefined at the turn of the last century to give greater emphasis to tackling inequalities in health (Graham (2004)). The World Health Organization (WHO) enshrined these goals in a Health for All strategy, which was aimed at reducing health inequalities within and between countries, where within-country targets focused on reducing inequalities between socioeconomic groups (World Health Organization (1999)). In England in 2001, a national target was set for narrowing health inequalities in infant mortality and life expectancy (Department of Health (2003)). The life expectancy target stipulated a reduction of at least 10 percent in the gap between the bottom quintile, based on health and deprivation indicators, and the population as a whole.

In 2008, an influential report by the Commission on Social Determinants of Health (World Health Organization (2008)) argued that social inequalities in power, money and resources operating throughout the life course, rather than individual unhealthy behaviors or access to health care, were instrumental in causing the observed social gradient in health.[1] The social gradient arises not only because of sharp differences in the health of the best- and worst-off in society, but also because those who are relatively disadvantaged, in terms of social position, have progressively worse health outcomes than those who occupy a higher rung in the social hierarchy. The subsequent Marmot review was tasked with supporting the development of a health inequalities strategy for England, including a monitoring framework for indicators and targets (Marmot Review Team (2010)). While not defining a specific health gain target, the report called for the development of an aspirational national health outcomes goal that included life and health expectancies achievable within the specified time scale and focused on the reduction of differences across the social gradient (pp. 166–7).

In this paper, we provide a tool to assist public policies in defining their health inequality strategy. The proposed framework is able to analyze the impact on life expectancy, by socioeconomic circumstances, of a hypothetical cause-of-death mortality reduction. Since social groups are affected differently by the causes of death, specific causes may need to be targeted in order to reduce inequalities. To gain such insight, we choose to use the multinomial logistic model developed by Alai *et al.* (2015a). This model readily quantifies the impact of cause-elimination, or shocks, on mortality metrics such as life expectancy. The model provides an intuitive framework for any combination of shocks on the vari-

[1] The social gradient refers to observed differences between socioeconomic groups for some indicators of interest, such as death rates, health status, etc. A higher socioeconomic gradient results in greater differences across socioeconomic groups for the chosen indicator.

ous considered causes and is readily extended to study the impact of these shocks across socioeconomic circumstances.

An important assumption of the cause-elimination mechanism in the model of Alai *et al.* (2015a) is that of extrinsic independence among the causes of death, a classic assumption in causal mortality models. Indeed, as *extrinsic* dependence among causal mortality rates is not objectively observable, notwithstanding the *intrinsic* dependence inherent in the nature of competing outcomes, an assumption needs to be set. The most widely used is the independence assumption, although theories and methods attempting to model the extrinsic dependence among causes have been proposed, including models incorporating individual risk factors (covariates, e.g., body mass index, blood pressure, smoking level; see, e.g., Girosi and King (2006); Manton (1986); Manton *et al.* (1991); Rosén (2006)); models incorporating individual unobserved risk factors, referred as frailties (Hougaard (1984); Manton *et al.* (1986); Vaupel and Yashin (1983)); models employing multiple cause-of-death data and thus providing a tool to investigate links between various causes (Mackenbach *et al.* (1999); Manton and Myers (1987); Manton and Poss (1979); Manton *et al.* (1976, 1980)); copulas (Carriere (1994); Dimitrova *et al.* (2013); Kaishev *et al.* (2007); Lo and Wilke (2010)); models using cointegrating techniques in order to capture the long-run equilibrium relationships that exist between different cause-specific mortality rates (Arnold and Sherris (2013, 2015, 2016)). Traditionally, the independence assumption developed by Chiang (1968) is applied; see, e.g., the United States decennial life tables by Anderson (1999); Bayo (1968); Curtin and Armstrong (1988); Greville *et al.* (1975). This approach is specified for causal forces of mortality. However, Alai *et al.* (2015a) recently proposed a different procedure, based on annual probabilities. In their approach, survival and death are competing outcomes and, therefore, treated similarly, as opposed to Chiang's model. We chose this approach because of the convenience and accessibility of working with annual probabilities within the multinomial framework.

Assessing the impact on life expectancy with a cause-elimination mechanism that incorporates socioeconomic variables is prudent, since the scale of the social gradient varies by cause; as different causes-of-death are linked to different risk factors, the mix of causes of death may differ across socioeconomic groups; and most importantly, it allows for scenario analysis to assess the causes of death most influential to gains in overall life expectancy, as well as reductions in life expectancy gaps.

The remainder of the paper is organized as follows. In Section 2, we introduce the data set and its characteristics regarding the causes of death and the socioeconomic categories. Section 3 outlines all relevant aspects of the methodology, including a description of the multinomial logistic model and the mechanism employed to shock mortality. After providing a subset of the model fit results in Section 4, we look at how to meet various policy targets using cause-elimination scenarios in Section 5. Section 6 concludes the paper.

2 Data Set

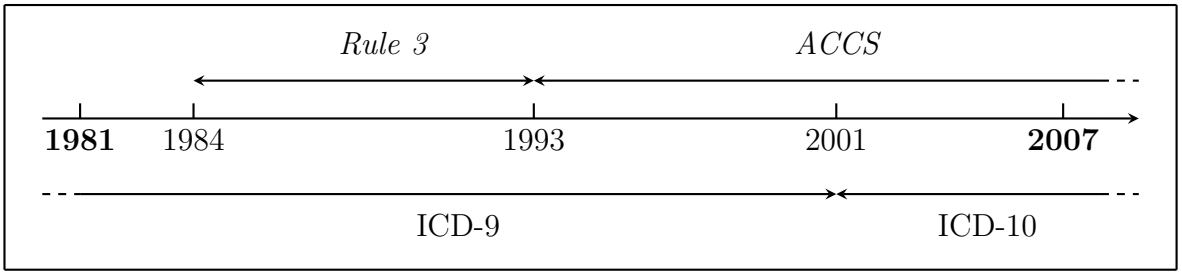
The cause-of-death mortality database was provided by the Office for National Statistics (ONS) in the United Kingdom. This database contains information on the gender, age, year, socioeconomic circumstances and cause of death of each registered death in England. The analyzed period is from 1981 to 2007, while the observed ages are grouped in five-year age classes, from age 25 to the open group 85+. The causes of death are grouped in six categories: diseases of the circulatory system, neoplasms, diseases of the respiratory system, diseases of the digestive system, external causes and a final "other" category (see

Appendix A for details). The corresponding central exposure figures were provided by ONS for the period 2001–2007, while midyear populations were estimated by Norman *et al.* (2008) for 1981–2000. In this section, we first detail the cause-of-death classification and the determination of the socioeconomic circumstances, before introducing some notations.

2.1 Cause-of-Death Classification

ONS provided counts of deaths aggregated up to three-digit codes of the International Classification of Diseases (ICD) developed by the World Health Organization (WHO). The ICD is used worldwide and is regularly updated to take into account progress in science and technology and to achieve more refined cause descriptions. Over the analyzed period, two different ICDs were used in England: ICD-9 until 2001 and ICD-10 from then until 2007. The classification of deaths has undergone two additional changes in England: *Rule 3* in 1984 and *ACCS* in 1993 (see Figure 1 for a summary). ONS used over the period 1984–1993 a guideline called Rule 3 in order to select the main cause of death (ONS (2015)). These guideline selection rules for the underlying cause of death were different from the rules established by ICD-9 (Rooney and Smith (2000)). Finally, in 1993, an Automated Cause Coding System (ACCS) was introduced; it is a computerized system for coding the causes of death from death registrations (for additional details, see Rooney and Devis (1995) and ONS (2015)).

Fig. 1: Death Classification Over the Period 1981–2007 in England



Naturally, as mentioned by Rooney and Smith (2000) and Villegas (2015), classification revisions affect cause-of-death mortality trends, so these revisions need to be carefully considered. Therefore, ONS developed comparability ratios in order to take into account the changes of classification from ICD-9 to ICD-10 (see details in Rooney *et al.* (2002)). We are directly using death rates adjusted by these comparability ratios (Bajekal *et al.* (2011); Villegas (2015)).

2.2 Socioeconomic Circumstances

The socioeconomic classification is established according to the Index of Multiple Deprivation of 2007 (IMD 2007). The IMD 2007 is the official measure of relative deprivation at the small living area level (Lower Layer Super Output Area, LSOA) in England. It is a weighted indicator based on the following factors: income, employment, health, education, barriers to housing and services, living environment and crime; for more details, see Noble *et al.* (2007). The IMD 2007 score allows classification of every LSOA, considered as homogeneous, into five equal-size deprivation categories or socioeconomic categories, with a

1 indicating the least deprived (highest quintile) and a 5 the most deprived (lowest quintile). To ease readability, we will also refer to quintile 1 as the "highest socioeconomic category" and quintile 5 as "lowest socioeconomic category," even if deprivation and the socioeconomic rank are not necessarily equivalent. The deprivation quintile allocation of an area is considered fixed over the complete period of observation; see Bajekal *et al.* (2013a) for details on this assumed stability and potential biases arising from it. Readers seeking additional details on the socioeconomic classification used should see Lu *et al.* (2014) and Villegas and Haberman (2014).

2.3 Notation

To convert deaths and central exposure to annual mortality rates and initial exposure, the following steps are applied. First, for each set of covariates, total deaths are calculated; these are used in order to establish initial exposure via the relationship

$$E^0(g,x,s,t) = E^c(g,x,s,t) + \frac{1}{2}D(g,x,s,t),$$

where $E^0(g,x,s,t)$, $E^c(g,x,s,t)$ and $D(g,x,s,t)$ denote initial exposure, central exposure and total deaths, respectively, for gender g, age x, socioeconomic circumstances s, and time t. Using initial exposure and cause-specific deaths, the cause-specific annual mortality rates are calculated via the relationship

$$q_i(g,x,s,t) = \frac{D_i(g,x,s,t)}{E^0(g,x,s,t)},$$

where $q_i(g,x,s,t)$ and $D_i(g,x,s,t)$ denote the annual mortality rate and deaths, respectively, for cause i. Lastly, survival is calculated via the relationship

$$p(g,x,s,t) = 1 - \sum_i q_i(g,x,s,t) = 1 - \frac{D(g,x,s,t)}{E^0(g,x,s,t)},$$

where $p(g,x,s,t)$ denotes the survival probability for a specific set of covariates.

3 Methodology

In this section, we provide the necessary modeling details. We outline the multinomial logistic model and the regression formula we chose in order to link the mortality rates to some selected covariates. We also detail the way shocks are applied to cause-specific death rates.

3.1 Multinomial Logistic Model

Multinomial logistic (or logit) regression techniques are useful in modeling probabilistic response variables for a competing categorical observations (see, e.g., Menard (2002) and Borooah (2002). These models were used to investigate cause-specific mortality over the entire age range in Alai *et al.* (2015a). Since this paper is extending the model introduced in Alai *et al.* (2015a) to a more comprehensive database, we briefly summarize their model and extend it to socioeconomic circumstances.

A multinomial logistic model is based, as indicated by its name, on a multinomial distribution: E realizations can be classified in $n+1$ different outputs, each with its own probability. In our study, E represents the initial exposure, while the $n+1$ different outputs represent n causes of death and survival. Indeed, among E exposed individuals, d_1 may die of cause 1, ..., d_n may die of cause n, and l may survive, with probability

$$\Pr[D_1 = d_1, \dots D_n = d_n, L = l] = \frac{E!}{d_1! \cdots d_n! l!} q_1^{d_1} \cdots q_n^{d_n} p^l,$$

where D_i denotes the random number of deaths from cause i, L denotes the subsequent survivors that complement the deaths, q_i describes the probability of death as a result of cause i, p the probability of survival, and

$$E = l + \sum_{k=1}^{n} d_k;$$

$$\sum_{k=1}^{n} q_k + p = 1.$$

For ease of notation, we omit the arguments defining the gender, age group, year and socioeconomic category. However, each variable should be understood with the additional arguments (g, x, s, t); e.g., $D_i(g, x, s, t)$ represents the random deaths from cause i, for gender g, age x, socioeconomic circumstances s and time t.

The multinomial *logit* model uses the *logit* transform of q_i in order to link the mortality and survival rates to a selection of covariates. Adopting survival as baseline category, we have

$$\log \frac{q_i}{p} = X\beta_i, \quad i = 1, \dots, n,$$

where X, named the design matrix, contains values of explanatory variables and β_i is the vector of regression parameters especially suited to cause i. The design matrix, X, may contain indicator or numerical variables for categorical or continuous covariates, respectively. The regression formula is the result of the product between the design matrix and the vector of regression parameters, which we outline in Section 3.2. Knowing the regression parameters and the design matrix, we find the probabilities of interest by applying the logistic function as follows:

$$q_i = \frac{\exp\{X\beta_i\}}{1 + \sum_k \exp\{X\beta_k\}}, \quad i = 1, \dots, n, \tag{1}$$

$$p = \frac{1}{1 + \sum_k \exp\{X\beta_k\}}. \tag{2}$$

Notice that the form of the survival probability, p, differs from the probabilities of death, q_i, since survival is designated as the baseline category.

3.2 Regression Formula

The explanatory variables used in the regression formula are a combination of time, age, gender and socioeconomic factors. The relationship is assumed linear as follows:

$$\begin{aligned}\ln \frac{q_i(g,x,s,t)}{p(g,x,s,t)} &= \beta_{0,i} + \beta_{1,g,i} + \beta_{2,x,i} + \beta_{3,s,i} + \beta_{4,i}t + \beta_{5,i}t^2 \\ &\quad + \beta_{6,g,x,i} + \beta_{7,g,s,i} + \beta_{8,g,i}t + \beta_{9,g,i}t^2 \\ &\quad + \beta_{10,x,s,i} + \beta_{11,x,i}t + \beta_{12,x,i}t^2 + \beta_{13,s,i}t + \beta_{14,s,i}t^2 \\ &\quad + \beta_{15,g,x,s,i} + \beta_{16,g,x,i}t + \beta_{17,g,x,i}t^2 + \beta_{18,g,s,i}t + \beta_{19,g,s,i}t^2.\end{aligned}$$

Time t is normalized to start with value 1, representative of year 1981. The regression parameters $\beta_{1,g,i}$, $\beta_{2,x,i}$ and $\beta_{3,s,i}$ are main gender, age group and socioeconomic circumstances specific, respectively. The regression parameters $\beta_{4,i}$ and $\beta_{5,i}$ describe the main linear and quadratic trends over time. Gender interaction parameters with age group, socioeconomic circumstances, linear and quadratic time are given by $\beta_{6,g,x,i}$, $\beta_{7,g,s,i}$, $\beta_{8,g,i}$ and $\beta_{9,g,i}$, respectively. Furthermore, $\beta_{10,x,s,i}$, $\beta_{11,x,i}$ and $\beta_{12,x,i}$ are the interaction parameters for age group with socioeconomic circumstances, linear and quadratic time, respectively. $\beta_{13,s,i}$ and $\beta_{14,s,i}$ capture the interaction between socioeconomic circumstances and linear and quadratic time; $\beta_{15,g,x,s,i}$, $\beta_{16,g,x,i}$, $\beta_{17,g,x,i}$, $\beta_{18,g,s,i}$ and $\beta_{19,g,s,i}$ are the corresponding gender interaction terms. These last five parameters ensure that men and women have completely distinct parameter sets in the model. Finally, $\beta_{0,i}$ denotes the intercept parameter designating the reference case, which is a man aged 25–29 in the least deprived socioeconomic circumstances.

3.3 Residual Life Expectancy

Since our data set is categorized by age groups, as opposed to single ages, we apply the abridged life table method to calculate life expectancy (see e.g. Chiang (1984)). A required input in this method is parameter a_x, where x designates the age group, that controls the relationship between central and crude mortality rates. This relationship has already been established in Section 2.3 and is consistent with uniformly setting $a_x \equiv 0.5$. It results in the following relationship between q, the crude, and m, the central mortality rate:

$$q = \frac{2m}{2+m}.$$

This relationship is widely used, the only exception being with infant mortality rates. As our database provides mortality rates from age 25, no additional assumption is required.

3.4 Cause-Specific Mortality Shocks

The *intrinsic* nature of the dependence among the causes of death is addressed by the multinomial model. If the relationship between the causes is stable and the causes themselves do not experience shocks, then this is sufficient for inferential as well as forecasting purposes. However, should one (or more) of the causes experience a positive shock, the other causes may adjust to this in an unpredictable way. The uncertainty surrounding the impact of causal shocks is due to *extrinsic* dependence. To address this, we use the independence assumption between the causes of death defined by Alai *et al.* (2015a).

Suppose we allow individual causes to receive shocks. Let ρ_i denote the marginal shock to cause i, and assume $0 \leq \rho_i$. The case $\rho_i = 0$ indicates the elimination of cause i.[2] Further, $0 < \rho_i < 1$ indicates a marginal reduction in cause i mortality, $\rho_i = 1$ indicates the absence of any marginal shock to cause i, and $\rho_i > 1$ indicates a marginal increase in cause i mortality. In the absence of any information of extrinsic dependence, mortality redistribution can be given by

$$q_i = \frac{\rho_i e^{X\beta_i}}{1 + \sum_k \rho_k e^{X\beta_k}}, \quad i = 1, \ldots, n, \tag{3}$$

$$p = \frac{1}{1 + \sum_k \rho_k e^{X\beta_k}}. \tag{4}$$

In the absence of shocks, Equations (3) and (4) reduce to Equations (1) and (2), respectively. Equations (3) and (4) represent a proportional reweighting of mortality, which is akin to assuming extrinsic *independence*. Besides, when cause j is eliminated within that framework, deaths from causes $i \neq j$ increase comparatively more, and survival increases comparatively less than previous findings using the independence assumption developed by Chiang (1968). Please see Alai *et al.* (2015a) for a more detailed comparison of the two approaches.

We note that in Equations (3) and (4), we have implicitly assumed that shocks are the same across ages and socioeconomic groups. However, population-level interventions to reduce health risk factors can inadvertently increase social inequalities in outcomes by disproportionately benefiting advantaged groups (Lorenc *et al.* (2013)). Thus, we could allow for socioeconomic-dependent shocks ρ_i^j, where i denotes the cause and j the socioeconomic category. However, since the focus of this paper is not to help inform public health policy to reduce inequalities in the social determinants of mortality, we assume as a starting point that any scenario affects the groups in the same way, so $\rho_i^1 = \ldots = \rho_i^5 = \rho_i$. We are indeed interested in modeling the *consequences* of achieving the stated mortality targets and/or the impact of complete cause elimination on life expectancy level and gap. Which practical policy steps are needed to get to this level of mortality reduction is not the question we are addressing in this paper. Besides, the Marmot Review suggests that policy makers adopt the approach of "proportional universalism" (Marmot Review Team (2010)). Using this approach, interventions are delivered to the whole population, with the "intensity" adjusted according to the needs of specific groups (for example, some groups may need more frequent help and advice). Such "upstream" actions are unlikely to eliminate the social gradient in health completely but may help to reduce the gap.

4 Model Fit

A selection of fitted values is presented in Figures 2 to 5. Figures 2 and 3 show observed log-mortality rates at age 65, with the fitted model, for males and females, respectively. The model appears to capture very well the observed trends of the data. It smooths out the noise considerably and achieves a well-defined distinction between each socioeconomic category—the top and bottom lines representing death rates for the most deprived and least deprived,

[2]The description of cause elimination in this paper is more faithfully represented by the idea of *ignoring* causes, rather than *eliminating* them, based on the definition of these terms introduced in Elandt-Johnson (1976). Henceforth, we continue to use the term "cause elimination," but this should not be confused with the definition outlined in Elandt-Johnson (1976). See Alai *et al.* (2015a) for details.

respectively. At the same time, the fitted model does not lose any of the essential patterns characterizing each cause of death. Similar observations are made regarding the fitted survival rates (Fig. 4) and life expectancies (Fig. 5): the model manages to smooth out the noise while capturing the underpinning trends for each cause of death and socioeconomic category.

Fig. 2: Observed and Fitted Values of Mortality Rates Over Time, Age Group 65–69, Males

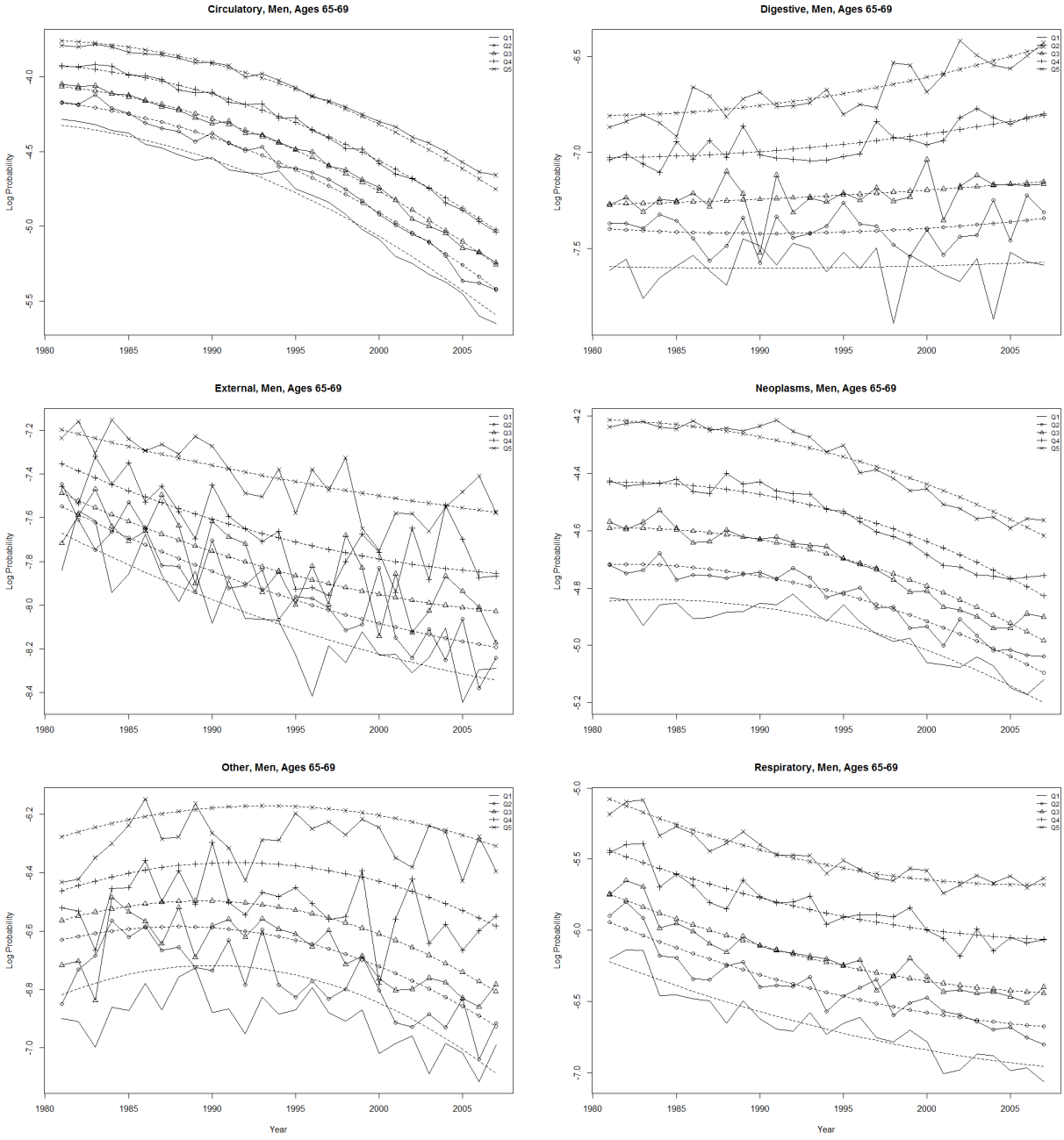


Fig. 3: Observed and Fitted Values of Mortality Rates Over Time, Age Group 65–69, Females

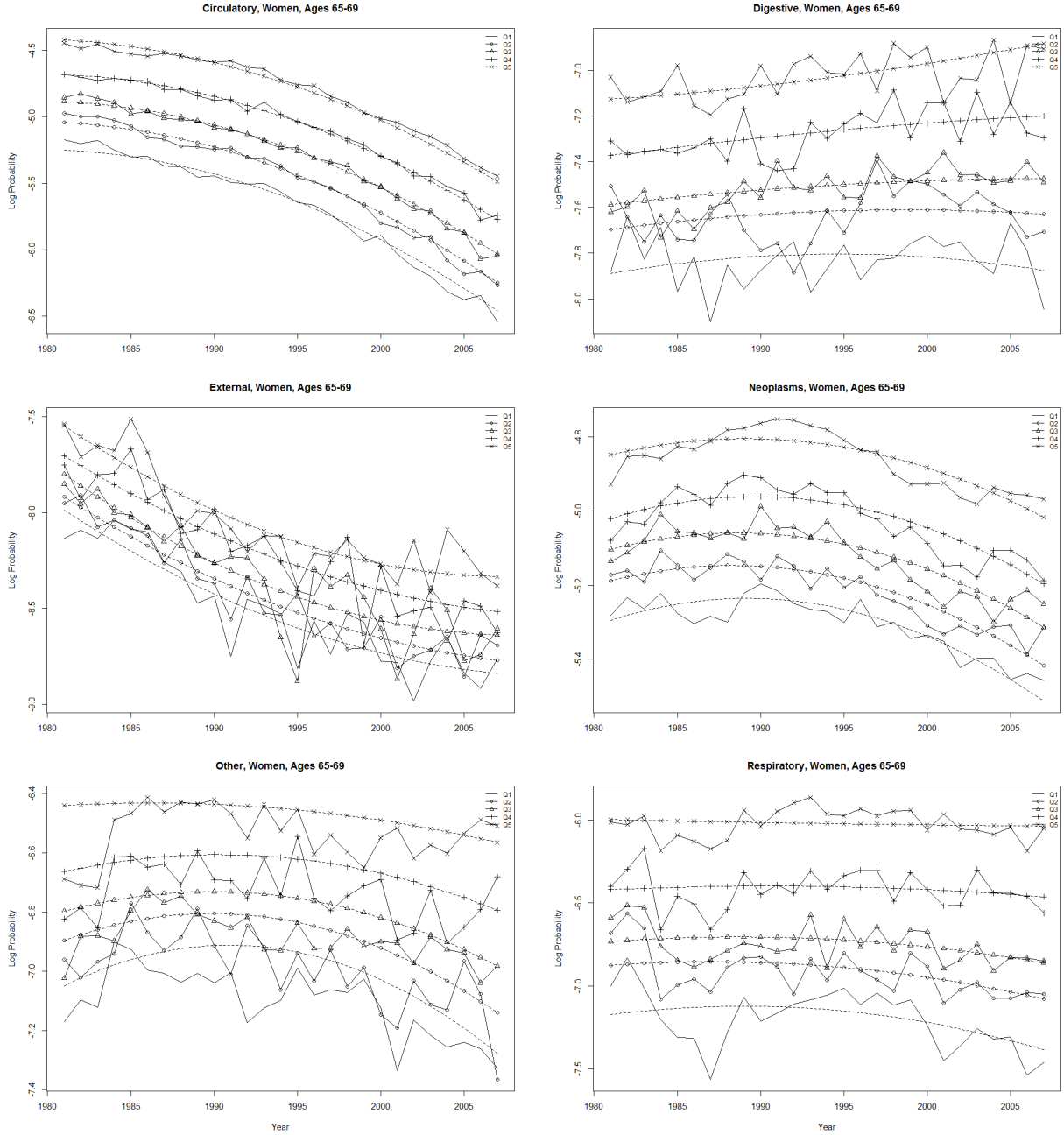


Fig. 4: Observed and Fitted Values of Survival Rates Over Time

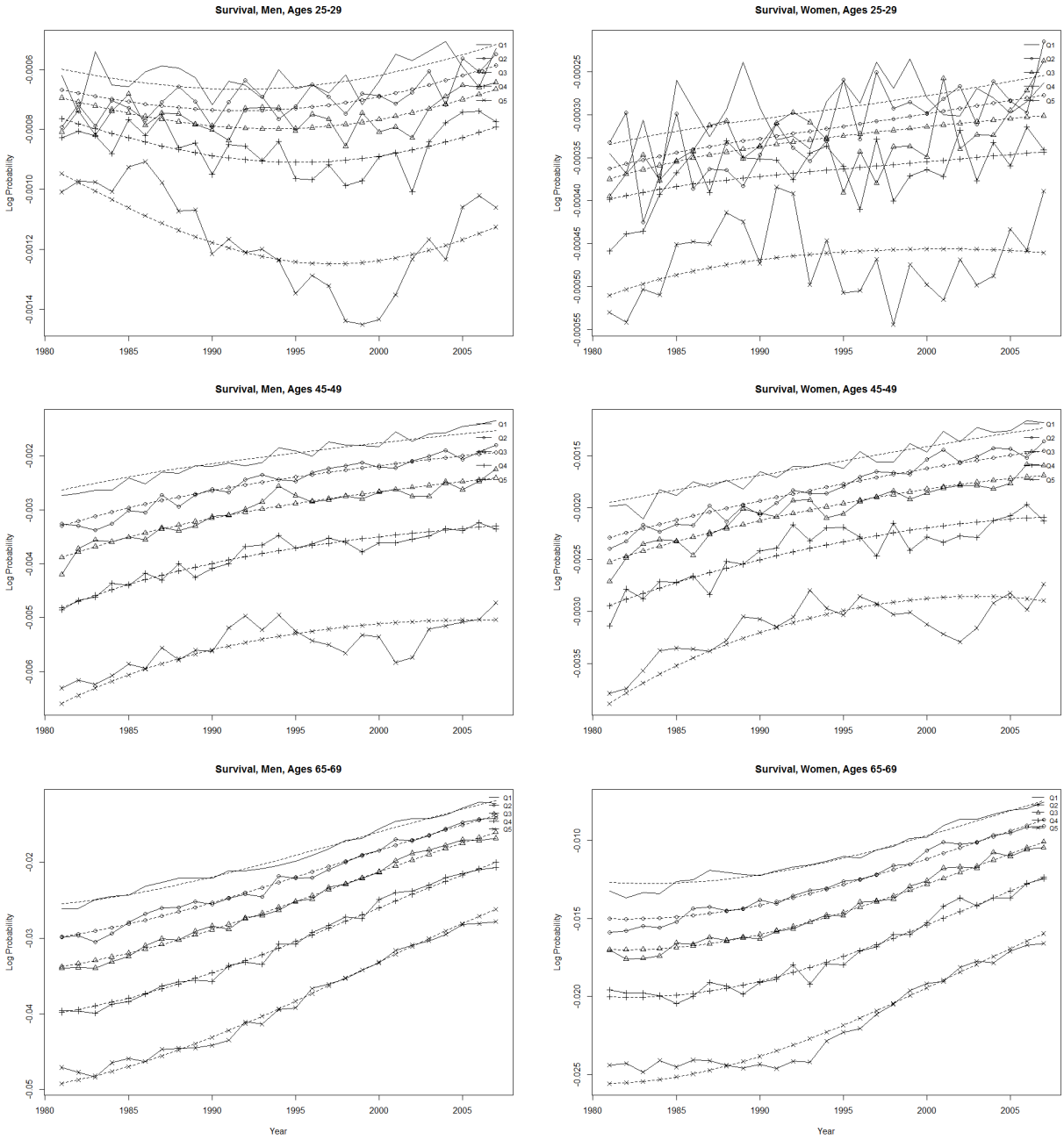


Fig. 5: Observed and Fitted Life Expectancies Over Time

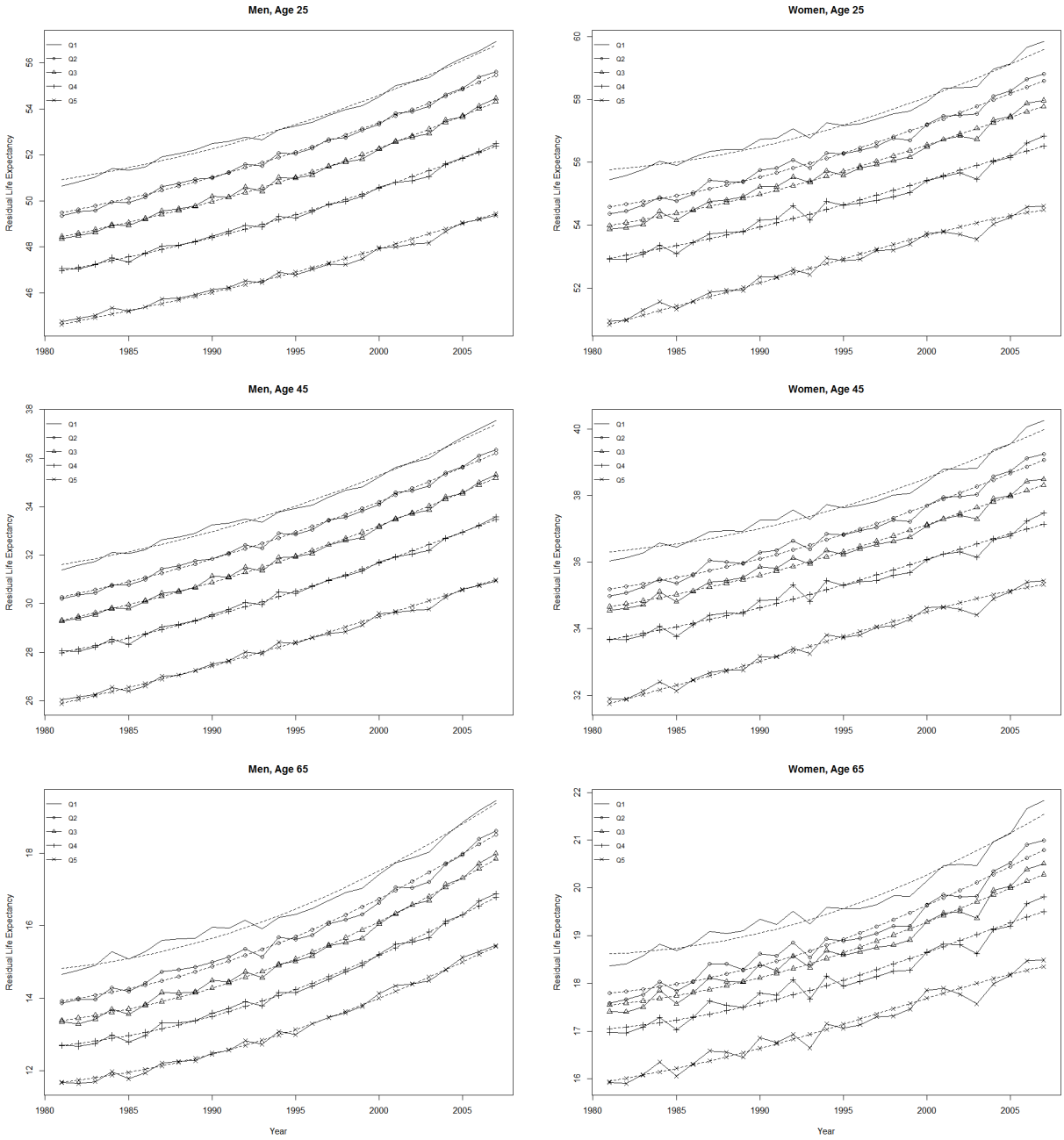


Table 1 provides an analysis of the effects in the model. It tests whether the inclusion of the effect represents an improvement in the model. A chi-square hypothesis test determines the significance of the effect. The table indicates that each effect significantly improves the model fit.

Table 1: Analysis of Effects

Effect	**DF**	**Wald Chi-Square**	**Pr > ChiSq**
gender (g)	6	345	<.0001
age (x)	72	579825	<.0001
imd07q (s)	24	1092	<.0001
period (t)	6	204	<.0001
period2 (t^2)	6	167	<.0001
gender*age	72	19678	<.0001
gender*imd07q	24	35	0.0724
gender*period	6	37	<.0001
gender*period2	6	31	<.0001
age*imd07q	288	70199	<.0001
age*period	72	5456	<.0001
age*period2	72	3546	<.0001
imd07q*period	24	183	<.0001
imd07q*period2	24	120	<.0001
gender*age*imd07q	288	2822	<.0001
gender*age*period	72	1622	<.0001
gender*age*period2	72	1103	<.0001
gender*imd07q*period	24	197	<.0001
gender*imd07q*period2	24	128	<.0001

5 Achieving Policy Aims: Scenarios

We now turn to quantifying the impact of some cause-elimination or cause-reduction scenarios on life expectancy, but most importantly on gaps in life expectancy existing between different socioeconomic categories. Causal-mortality shocks as described in Section 3.4 are applied, first by eliminating one cause of death at a time (Section 5.1), then by combining several cause-specific mortality shocks in order to reflect a policy target defined by the World Health Organization (Section 5.2). Finally, an optimization approach is developed in Section 5.3 to increase life expectancy and reduce inequalities simultaneously.

5.1 Scenario 1: Eliminating One Cause of Death at a Time

We successively apply six different mortality shocks: we eliminate each of the six causes of death in turn; for example, we set the ρ parameter of diseases of the circulatory system to 0 while keeping all the other ρ parameters equal to 1 in Equations 3 and 4. Table 2 shows the gain in life expectancy for men aged 65 in 2007 by socioeconomic circumstances when deaths from each of the causes are eliminated. We see, for example, that the elimination of circulatory diseases and neoplasms would produce the greatest gains, respectively, for men

aged 65. It is also interesting to note that, in this case, the relative gains always favor the lower socioeconomic groups; deaths by digestive and respiratory causes even favor lower socioeconomic groups in absolute terms. Table 3 shows the same results for women aged 65 in 2007.

Table 2: Residual Life Expectancy for Men Aged 65, 2007

Residual Life Expectancy	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**
Fitted	19.38	18.51	17.84	16.78	15.42
Fitted {–circulatory}	23.93	22.92	22.24	20.99	19.37
Absolute gain	4.55	4.41	4.40	4.21	3.95
Relative gain	123%	124%	125%	125%	126%
Fitted {–digestive}	19.7	18.84	18.17	17.13	15.79
Absolute gain	0.32	0.33	0.33	0.35	0.37
Relative gain	102%	102%	102%	102%	102%
Fitted {–external}	19.53	18.66	17.99	16.93	15.56
Absolute gain	0.15	0.15	0.15	0.15	0.14
Relative gain	101%	101%	101%	101%	101%
Fitted {–neoplasms}	22.5	21.52	20.78	19.63	18.20
Absolute gain	3.12	3.01	2.94	2.85	2.78
Relative gain	116%	116%	116%	117%	118%
Fitted {–other}	20.54	19.61	18.87	17.75	16.29
Absolute gain	1.16	1.10	1.03	0.97	0.87
Relative gain	106%	106%	106%	106%	106%
Fitted {–respiratory}	20.8	19.95	19.33	18.36	17.09
Absolute gain	1.42	1.44	1.49	1.58	1.67
Relative gain	107%	108%	108%	109%	111%

Table 3: Residual Life Expectancy for Women Aged 65, 2007

Residual Life Expectancy	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**
Fitted	21.54	20.79	20.28	19.5	18.35
Fitted {–circulatory}	26.21	25.27	24.76	23.87	22.42
Absolute gain	4.67	4.48	4.48	4.37	4.07
Relative gain	122%	122%	122%	122%	122%
Fitted {–digestive}	21.94	21.21	20.71	19.95	18.82
Absolute gain	0.40	0.42	0.43	0.45	0.47
Relative gain	102%	102%	102%	102%	103%
Fitted {–external}	21.71	20.96	20.45	19.66	18.51
Absolute gain	0.17	0.17	0.17	0.16	0.16
Relative gain	101%	101%	101%	101%	101%
Fitted {–neoplasms}	23.89	23.10	22.61	21.82	20.70
Absolute gain	2.35	2.31	2.33	2.32	2.35
Relative gain	111%	111%	111%	112%	113%
Fitted {–other}	23.46	22.63	22.05	21.14	19.85
Absolute gain	1.92	1.84	1.77	1.64	1.50
Relative gain	109%	109%	109%	108%	108%
Fitted {–respiratory}	23.02	22.30	21.85	21.22	20.21
Absolute gain	1.48	1.51	1.57	1.72	1.86
Relative gain	107%	107%	108%	109%	110%

Naturally, the elimination of the deaths from a specific cause increases life expectancy in each socioeconomic category. However, this will not necessarily reduce the social gaps. Figure 6 presents the time evolution of the gap in life expectancy at ages 25, 45 and 65 between the lowest and highest socioeconomic groups for men and women. The black line labeled "Observed" reproduces the gap that was prevailing from 1981 to 2007. It is interesting to note that the gap for women decreased until approximately the mid-1990s and started to increase thereafter, while for men the gap at ages 25, 45 and 65 increased from the mid-1980s, end of the 1980s and over the whole period, respectively. Each plot also contains the gap in life expectancy that would result if a specific cause of death were eliminated. For example, the line labeled "-circulatory" would reflect the gap in life expectancy between the lowest and highest socioeconomic categories if deaths due to the circulatory system were eliminated in 1981.

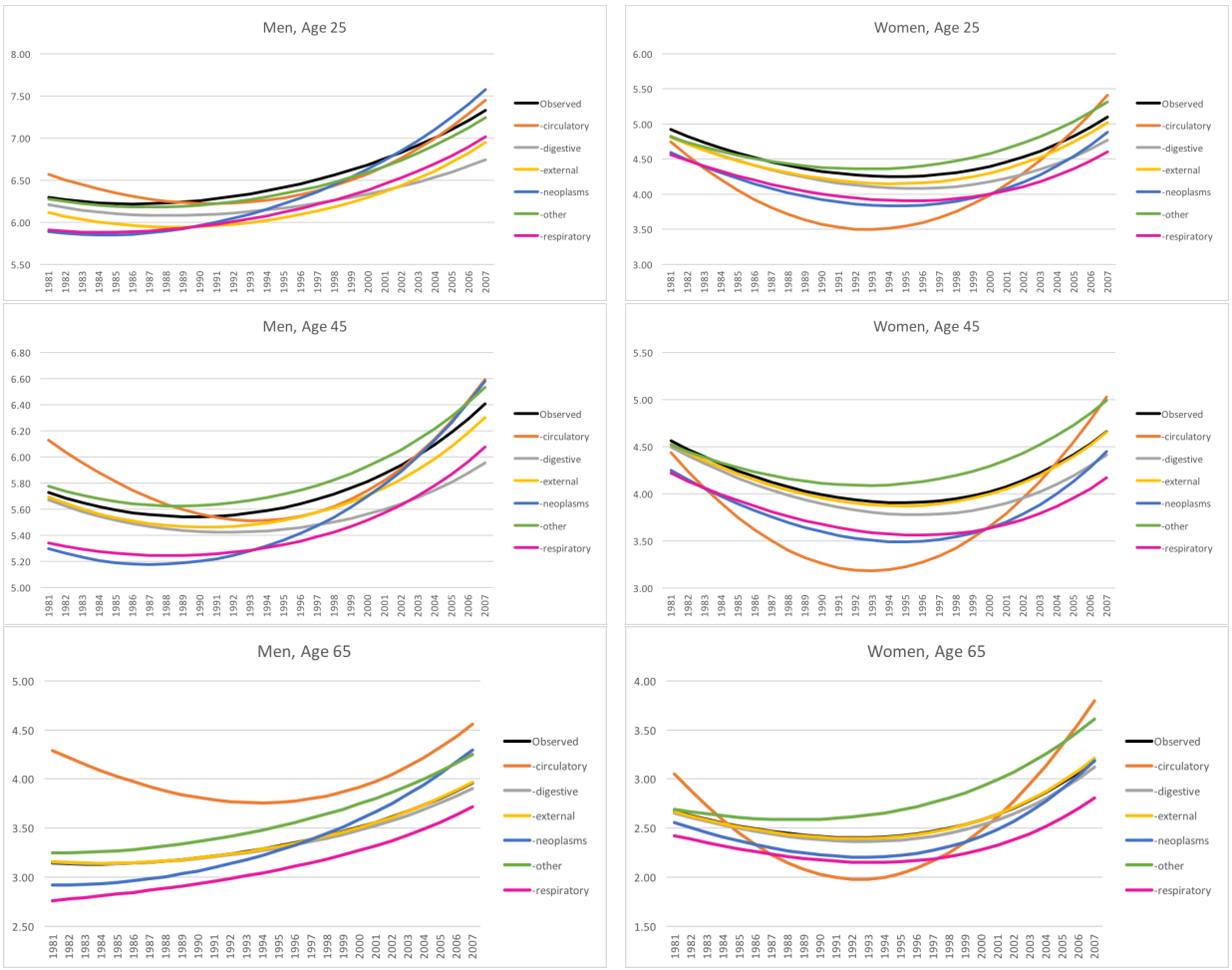
Several observations can be made about the impact of eliminating a cause in 1981 on the gap in life expectancy until 2007:

- Eliminating diseases of the circulatory system increases the inequality gap (compared with the observed all-cause mortality) for both men and women, especially at age 65 as well as at the beginning and end of the period of observation. This is in line with observations made in Nordic countries (Brønnum-Hansen and Baadsgaard (2012)).

- Eliminating neoplasms results in an increase in the gap for men in recent years, while it decreases the gap for women.

- The reduction in the gap following the elimination of neoplasms for women is less important nowadays than 25 years ago.

Fig. 6: Impacts of Cause Elimination on the Gap in Life Expectancy Between the Highest and Lowest Socioeconomic Categories



- Eliminating the "other causes" category increases the gap at ages 25, 45 and 65 for women, while the gap increases only at ages 45 and 65 for men.
- Eliminating digestive diseases decreases the gap in life expectancy at ages 25, 45 and 65 for both genders, and the reduction in the gap increases over time.
- Eliminating respiratory diseases decreases the gap at ages 25, 45 and 65 for both genders, the reduction being relatively constant over time.
- Eliminating the external causes of death decreases the gap at age 25 (both genders) and age 45 (men only), the impact being much smaller for women than for men.

The increase in the gap following the elimination of neoplasms or the diseases of the circulatory system may seem counterintuitive, since death rates for any cause are higher for the lowest socioeconomic category, and thus, eliminating a cause will save more lives in the lowest quintile. However, these saved lives have a higher probability to die of the remaining causes of death. Therefore, if the causes not deleted have a relatively greater socioeconomic gradient (e.g., respiratory diseases) than the cause deleted (e.g., circulatory diseases), inequality will increase. But if it were the other way around—if the differences in death rates between socioeconomic groups for the remaining causes were not as important as for the deleted cause—then inequality would decrease. Finally, eliminating a cause (e.g., external causes at age 65) would make no difference to the gap, if the gradient for this cause were similar to that of all causes. Therefore, the relative difference in the inequality gradient between the deleted and remaining competing causes will determine the magnitude and the direction of the inequality change from the baseline (here no cause elimination).

Another important remark is related to the impact of cause elimination on the gap across age groups. Figure 6 shows that the elimination of the diseases of the digestive system would reduce the gap importantly at ages 25 and 45 (especially nowadays), while it would not affect much the gap at age 65. The reason is simple: deaths due to digestive diseases occur mainly around age 45 and thus by age 65 are relatively insignificant.

5.2 Scenario 2: Meeting the WHO 25×25 Target

The WHO Global Burden of Disease initiative identified reduction in the health burden of noncommunicable diseases (NCDs) as a major issue for sustainable development. In response, the UN General Assembly signed a declaration committing member states to the prevention and control of NCDs (United Nations (2011)). Countries agreed to adopt an overarching target of reducing premature mortality (ages between 30 and 70) from the four main NCDs (cardiovascular diseases, chronic respiratory diseases, cancers and diabetes) by 25% relative to their 2010 levels by 2025 (referred to as the 25×25 target).

Since a key feature of the cause-elimination approach introduced in this paper is that it allows the consideration of simultaneous shocks to different causes, the WHO 25×25 target can be applied and its impact on life expectancy and social inequalities analyzed. However, the ICD classification used in our database does not exactly coincide with the required classification to analyze the WHO target. Indeed, we cannot analyze diabetes as a separate category (which represent around 1% of deaths in England), and we are additionally including acute respiratory conditions such as pneumonia in the respiratory category. Nevertheless, as the implied differences are minimal, applying the 25×25 target by reducing mortality from circulatory diseases, respiratory diseases and neoplasms by 25% still provides a very good estimation of the potential impacts of the WHO target on life

expectancy. Besides, this 25% decrease is applied across all ages, even if the 25×25 target focuses only on premature mortality. Indeed, we assume that actions to prevent premature deaths from NCDs will act at successively older ages. However, this assumption can easily be changed, the model allowing different mortality reductions across age groups.

Within our model framework, the 25×25 target is equivalent to setting in Equations 3 and 4, $\rho_i = 0.75$ for circulatory diseases, respiratory diseases and neoplasms while maintaining $\rho_i = 1$ for the other three groups of causes. A second, more aggressive target also is analyzed: since the 25×25 target is planned over 15 years (2010–2025), this would correspond to a 40% decline in 25 years (1981–2007). Thus, we also present results for a 40% mortality decrease for circulatory diseases, respiratory diseases and neoplasms ("40 target"). This is equivalent to setting $\rho_i = 0.6$ for these three groups of causes of death. Tables 4 and 5 present, for men and women, respectively, the remaining life expectancies at ages 25, 45 and 65 for each deprivation quintile under each of the two WHO target scenarios. We see that for both men and women and at at all ages and deprivation quintiles, achieving this reduction in cause-specific mortality results in a significant increase in life expectancy. For instance, for men aged 25, life expectancy would increase from an average across quintiles of 53.67 years in 2007 to 55.99 years under the the 25×25 target. However, it is worth noting that the achievement of the WHO targets might result in an increase in life expectancy inequalities at some ages. For example, for men aged 65, the 40 target would result in a 0.12-year increase in the difference in life expectancy between the least and most deprived quintiles of England.

Table 4: Residual Life Expectancy for Men in 2007 Under WHO Scenarios

Age 25							
Scenario	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Average**	**Q1–Q5**
Fitted 2007	56.77	55.47	54.31	52.38	49.44	53.67	7.33
25×25 Target	59.08	57.76	56.62	54.71	51.79	55.99	7.29
40 Target	60.91	59.57	56.62	56.52	53.60	57.44	7.30
Age 45							
Scenario	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Average**	**Q1–Q5**
Fitted 2007	37.39	36.21	35.18	33.47	30.99	34.65	6.40
25×25 Target	39.67	38.57	37.46	35.77	33.31	36.94	6.36
40 Target	41.48	40.26	39.26	37.57	35.11	38.74	6.37
Age 65							
Scenario	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Average**	**Q1–Q5**
Fitted 2007	19.38	18.51	17.84	16.78	15.42	17.59	1.79
25×25 Target	21.41	20.49	19.82	18.73	17.33	19.56	1.85
40 Target	23.06	22.10	21.42	20.31	18.87	21.15	1.91

5.3 Scenario 3: Optimal National Targets to Increase Life Expectancy and Reduce Inequalities

This raises the question of which scenario of cause elimination would help to close the life expectancy gap while achieving the greatest overall increase in life expectancy across the society. We can answer such a policy question using a multi-objective optimization approach whereby we seek to find (under some constraints) the scenario of cause-specific

Table 5: Residual Life Expectancy for Women in 2007 Under WHO Scenarios

Age 25							
Scenario	**Q1**	**Q2**	**Q3**	**Q4**	Q5	**Average**	**Q1–Q5**
Fitted 2007	59.58	58.60	57.77	56.50	54.49	57.39	5.09
25×25 Target	61.70	60.71	59.93	58.73	56.77	59.57	4.93
40 Target	63.34	62.32	59.93	60.43	58.50	60.90	4.84
Age 45							
Scenario	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Average**	**Q1–Q5**
Fitted 2007	40.00	39.07	38.32	37.14	35.33	37.97	4.66
25×25 Target	42.08	41.13	40.43	39.31	37.56	40.10	4.51
40 Target	43.68	42.72	42.05	40.98	39.25	41.74	4.43
Age 65							
Scenario	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Average**	**Q1–Q5**
Fitted 2007	21.54	20.79	20.28	19.50	18.35	20.09	3.18
25×25 Target	23.40	22.62	22.13	21.37	20.22	21.95	3.18
40 Target	24.87	24.05	23.58	22.84	21.68	23.40	3.19

reduction shocks that simultaneously maximizes the overall gain in life expectancy and minimizes the gap in life expectancy between the least and most deprived quintiles.

Formally, let $e_{x,g,s,t}(\boldsymbol{\rho})$ denote the remaining life expectancy at age x for gender g, socioeconomic circumstance s and time t, given a cause-reduction scenario $\boldsymbol{\rho} = (\rho_1, \ldots, \rho_n)$. Then, for fixed age x, gender g and time t, the optimization problem can be formulated as follows:

$$\min \quad \{f_1(\boldsymbol{\rho}) = -\tfrac{1}{5}\textstyle\sum_s e_{x,g,s,t}(\boldsymbol{\rho}), \quad f_2(\boldsymbol{\rho}) = e_{x,g,Q1,t}(\boldsymbol{\rho}) - e_{x,g,Q5,t}(\boldsymbol{\rho})\}, \tag{5}$$

s.t.

$$\sum_{i=1}^{n} (1 - \rho_i) \leq K; \tag{6}$$

$$\underline{\rho} \leq \rho_i \leq 1 \qquad i = 1, \ldots, n, \tag{7}$$

where the first objective f_1 is equivalent to maximizing the average life expectancy across deprivation quintiles and the second objective f_2 minimizes the absolute difference in life expectancy between the least deprived quintile (Q1) and the most deprived quintile (Q5). In the above optimization problem, Equation (6) is a total shock constraint, indicating that a total shock of K is allowed across all causes, while constraint (7) indicates that marginal increases are not allowed and that a cause can be only reduced up to a level $\underline{\rho}$.

We now use the optimization problem to evaluate whether the WHO 25×25 and WHO 40 targets are "optimal." To do so, we have solved the optimization problem with

- $\underline{\rho} = 1 - 0.25 = 0.75$ and $K = 3 \times 0.25 = 0.75$ for the WHO 25×25 target, and
- $\underline{\rho} = 1 - 0.4 = 0.6$ and $K = 3 \times 0.4 = 1.2$ for the WHO 40 target.

By setting $\underline{\rho} = 0.75$ for the WHO 25×25 target, we impose a maximum mortality decrease of 25% for each of the six causes. Besides, by imposing $K = 0.75$, we require that the cumulative mortality decrease across the causes is of maximum 75%. This maximum cumulative mortality decrease can be achieved through different combinations: (1) reducing only one cause-specific mortality rate by 75% and keeping all the other mortality rates

constant; (2) reducing one cause-specific mortality rate by 50% and another one by 25%; (3) reducing each of the six analyzed cause-of-death mortality rates by 10% (since the total needs to be equivalent to or smaller than 75%); etc. Thus, as in the WHO 25×25 target, we allow a mortality decrease of maximum 25% per cause and a total shock of maximum 75%. However, we do not pre-specify the causes of death that need to be reduced. The optimization procedure will determine, among all the possible combinations, which one corresponds to the greatest gain in life expectancy and largest reduction in social inequalities. The same approach is applied for the WHO 40 target.

Figure 7 presents for each sex and ages 25, 45 and 65 the Pareto front for both WHO target scenarios. These plots depict on each axis each of the two objectives, indicating what is the maximum average life expectancy, f_1, that can be achieved for a given gap in life expectancy, f_2. Figures 8 and 9 plot for the WHO 25×25 and the WHO 40 targets, respectively, the cause-specific shocks, ρ_i, that would need to be applied to achieve a given gap in life expectancy. For example, for men aged 25 and the 25×25 target scenario, Figure 8 shows that to achieve a gap in life expectancy of 7.1 years, diseases of the respiratory system need to be reduced by 20% ($\rho_{respiratory} = 0.8$), neoplasms need to be reduced by 5% ($\rho_{neoplasms} = 0.95$), and digestive and circulatory diseases need to be reduced by 25% ($\rho_{digestive} = \rho_{circulatory} = 0.75$). In parallel, an inequality gap of 7.1 years for men aged 25 corresponds to an average life expectancy across social groups of about 55.3, as described by Figure 7.

From these three figures we note the following:

- The conflicting nature of the objectives becomes clear with higher average gains in life expectancy coming at the cost of a wider life expectancy gap. For instance, under the WHO 40 policy scenario, the minimum life expectancy gap for men aged 25 is 6.82 years, which corresponds to the lowest average life expectancy (54.6 years), and the maximum average life expectancy is 57.81, which corresponds to the highest life expectancy gap (7.3 years; Figure 7).

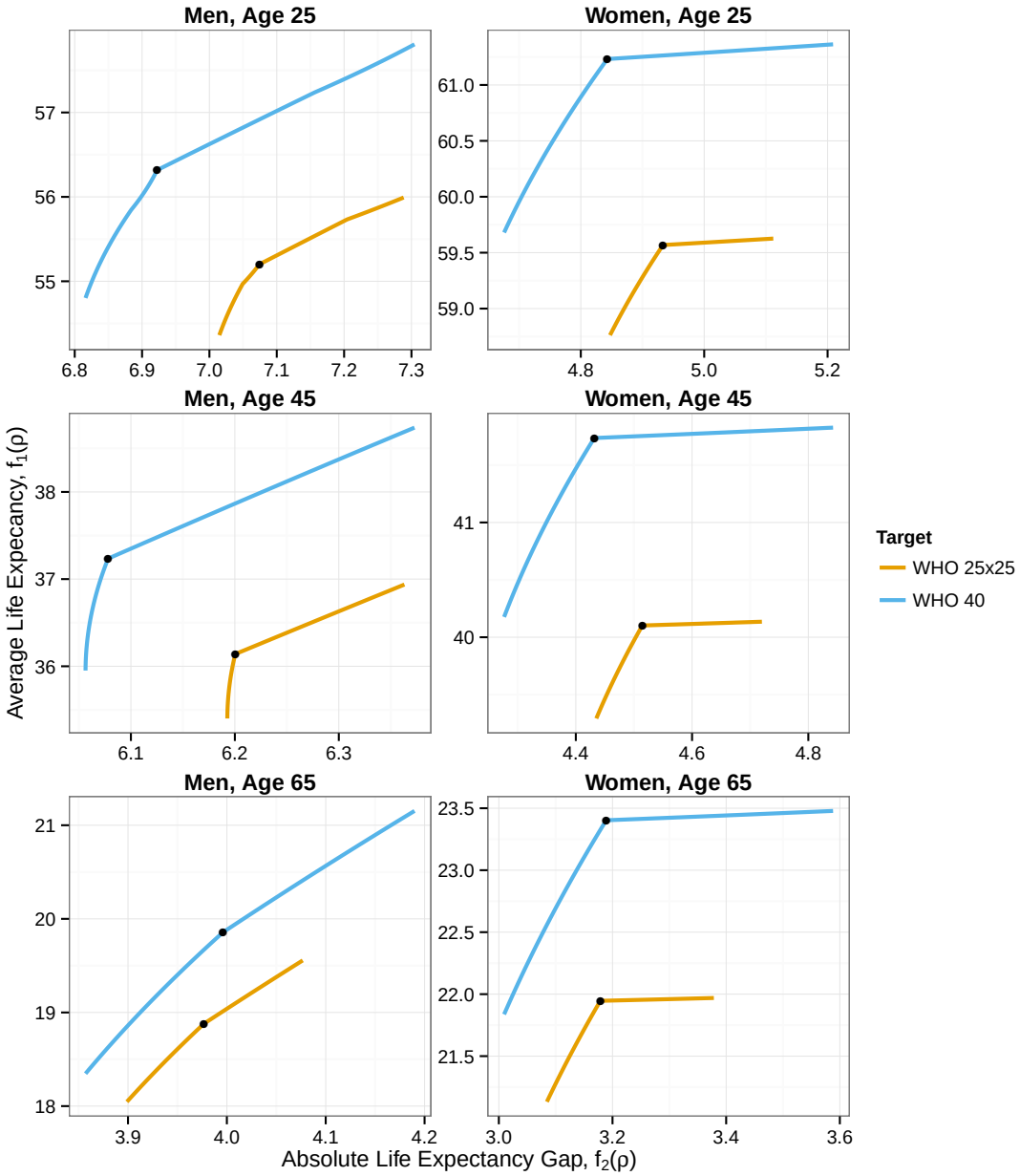
- While for women the optimal shocks are independent of age, for men the optimal policy is age dependent. (Figures 8 and 9 differ by age for men, as opposed to women.)

- In particular, for women of all ages, if the aim is to increase life expectancy independently of life expectancy gap (top-right end of plots in Figure 7), then the optimal policy target should be to reduce mortality from circulatory diseases, neoplasms and other causes of death (right end of plots in Figures 8 and 9). By contrast, for men of all ages, the maximum increase in life expectancy is achieved by targeting mortality from circulatory diseases, neoplasms and respiratory diseases.

- For women of all ages, if the aim is to decrease the life expectancy gap independently of increase in life expectancy (bottom-left end of plots in Figure 7), then the optimal policy target should be to reduce mortality from neoplasms and from respiratory and digestive diseases (left end of plots in Figures 8 and 9). By contrast, for men of all ages, the minimum gap in life expectancy is achieved by targeting digestive, respiratory and external causes. Circulatory diseases also need to be targeted for men aged 45.

Of main interest in multi-objective optimization problems are the so-called knees of the Pareto front (Branke *et al.*, 2004)—the solutions around which a marginal increase in life

Fig. 7: Pareto Front of the Two Optimal WHO Targets



Note: Black dots represent the "knees" of the Pareto front.

Fig. 8: Optimal Policies for WHO 25×25 Target

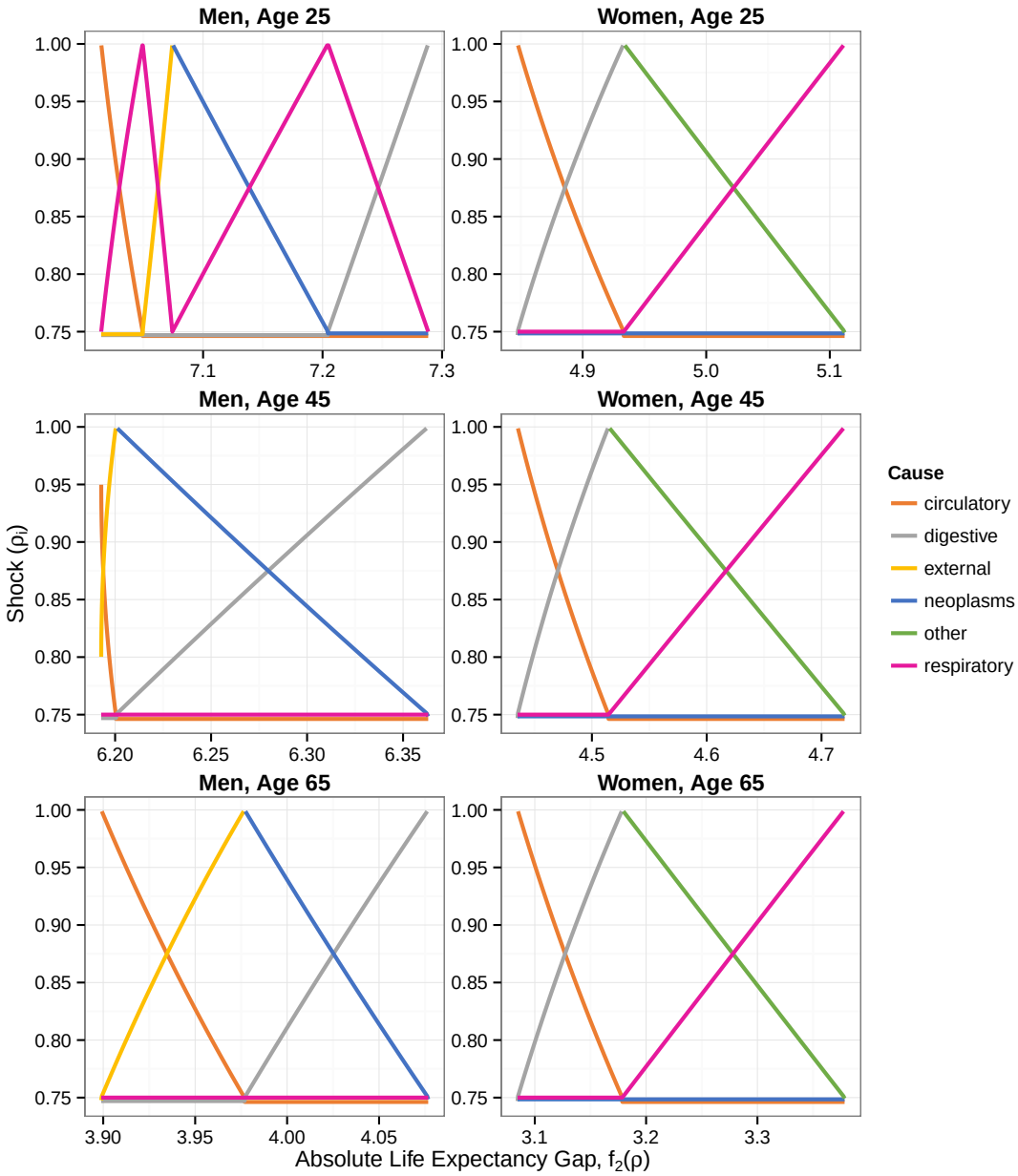
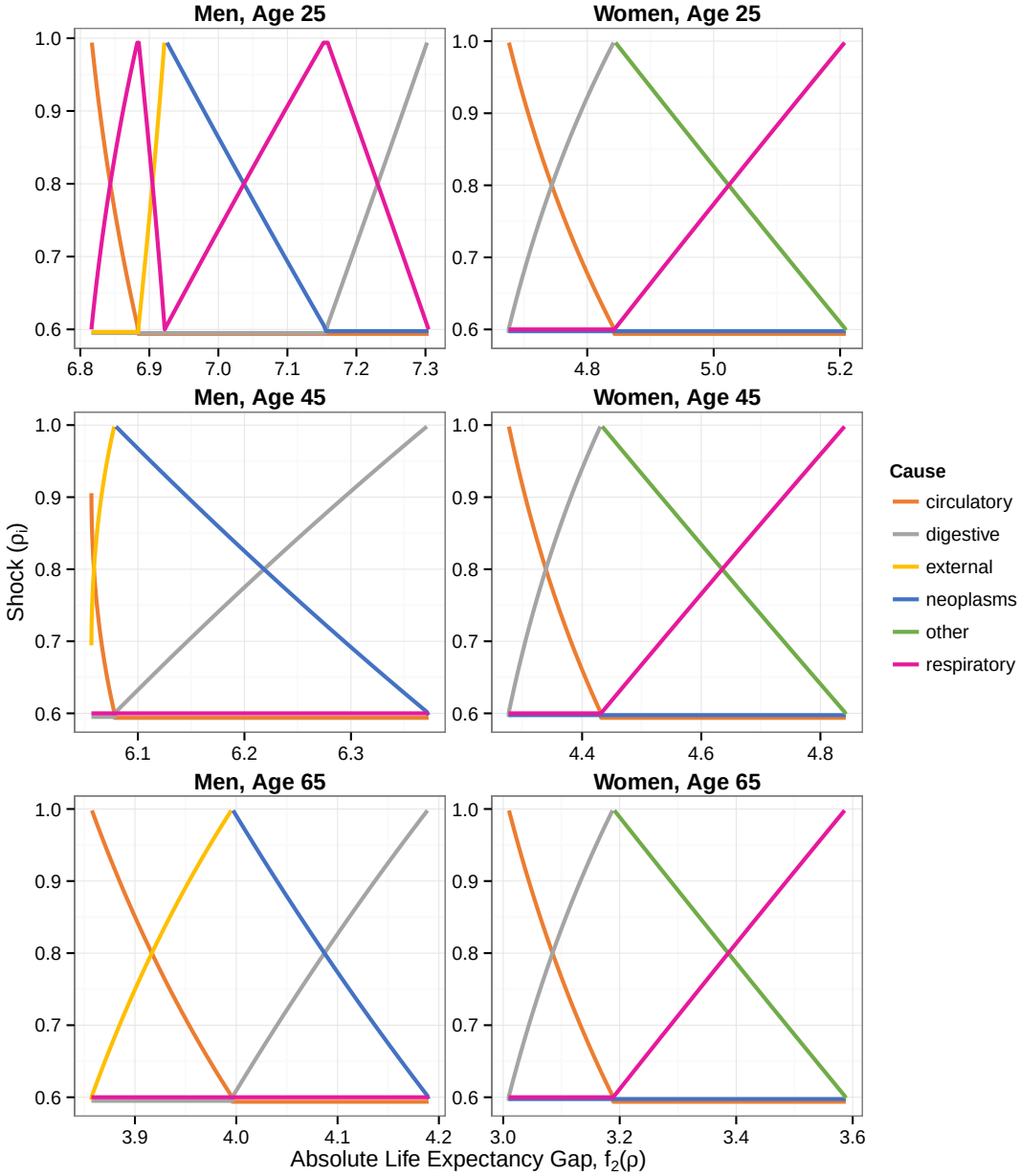


Fig. 9: Optimal Policies for WHO 40 Target



expectancy would result in a significant increase in life expectancy inequalities or where a marginal decrease in inequalities will result in significant reduction in the average life expectancy. In the absence of any additional information, such solutions are likely to be the ones preferred by the policy maker.

For women of all ages, the "knee" solutions correspond to the policy scenario targeting mortality from circulatory diseases, neoplasms and respiratory diseases. Interestingly, this coincides with the original WHO targets, suggesting that the WHO policy is optimal for women. By contrast, for men of all ages, the "knee" solutions correspond to the policy scenario targeting mortality from circulatory, digestive and respiratory diseases, which does not coincide with the original WHO targets. Hence, for men, a possibly more appropriate policy would be targeting digestive diseases, which are one of the main contributors to socioeconomic inequalities, as opposed to neoplasms.

6 Conclusion

This paper provides the basis to assist government bodies in implementing well-informed strategies aimed at reducing social inequalities. This is especially relevant today in England and around the world, given that addressing such inequalities has become a key focus of public policy. The proposed approach uses cause-specific mortality data in order to gain insight into differences in life expectancies by deprivation categories. We investigate the relationship between socioeconomic circumstances and causal mortality on a unique data set obtained from the UK Office for National Statistics and apply a model that has the ability to incorporate any combination of cause-specific mortality shocks. This investigation is performed in two steps. First, the model developed by Alai *et al.* (2015a) is extended to allow for socioeconomic covariates. Second, an optimization procedure is developed that identifies the cause-specific reduction scenario that simultaneously maximizes the overall gain in life expectancy while minimizing social inequalities.

The results presented in this paper may have important consequences in forming public policy. First, it is well known that the decline in heart disease mortality was a major contributor to increases in life expectancy over the past 25 years for men and women across all socioeconomic groups (Bajekal *et al.* (2013b)). However, the elimination or reduction of such causes of death contributes, according to our results, to an increase in inequalities by socioeconomic circumstances. Public policy exclusively targeting cardiovascular diseases may, therefore, be contradictory to the aim of reducing the social gradient in life expectancy.

Second, insight into the latest time trends is a crucial factor for determining policy aimed at a reduction in inequalities. Indeed, by 2007, the picture has significantly changed in terms of identifying the optimal cause of death to target in order to reduce life expectancy gaps. For example, eliminating neoplasms would have reduced the gap in the 1980s, while doing so more recently would have actually increased it.

Third, if the aim is to reduce the inequalities for all age groups and both genders, then, based on the situation in 2007, respiratory diseases need to be targeted for elimination. Since these deaths are largely related to smoking habits, prevention campaigns and other forms of dissuasion should be implemented. Public policy could also focus on targeting digestive diseases (e.g., liver cirrhosis) and external causes of death (accidents) in order to reduce the inequality gaps at younger ages, particularly for men.

Fourth, the targets set by WHO result in an increase in life expectancy for all groups, but also result in increasing inequalities for men. The developed optimization procedure indicates that it would be better for the male population to target digestive diseases instead

of neoplasms; this would achieve an optimal balance between increasing life expectancy and reducing the inequality gap.

Finally, it is worth highlighting that the way in which dependence is treated plays a very important role in our model framework. Including different dependence assumptions in our model would provide important and complementary information. For example, interesting properties of various multivariate distributions have recently been formulated that may be applicable to the study of cause-of-death dependence. In the work of Alai *et al.* (2013), a multivariate gamma distribution based on a common shock model was applied to lifetime data and later generalized to include Tweedie distributions in Alai *et al.* (2015b, 2016b). If the link between different causes is established via shared risk factors, a common shock model may be appropriate. Furthermore, a multivariate Pareto distribution with a parametric form of dependence was shown to be well estimated via the distribution of the minimum value of its marginals (see Alai *et al.* (2016a)); this may render it suitable for cause-of-death lifetime data, where only the minimum observation is available.

Naturally, any study on mortality reduction necessarily leads to the very interesting and controversial discussion of the biological processes of aging, considered by many as the greatest underlying risk factor affecting mortality (see, e.g., Hayflick (2004)). While some experts, such as De Grey (2006), expect huge mortality improvements for the coming century, others are more pessimistic. They argue that the human body was not designed for long-term use, so humans cannot biologically live much longer unless we manage, in the future, to alter our basic biology (see, e.g., Olshansky *et al.* (2002)). The latter case implies that if we managed to reduce current health inequalities by attacking one disease at a time, a new set of health inequalities would emerge, as the body is biologically aging and thus malfunctioning. Inequalities will then never disappear. Such questions are naturally important to study but are well beyond the scope of this paper.

To conclude, the optimization procedure and corresponding results introduced in Section 5.3 are presented for illustrative purposes. This work aims to show the wide variety of analyses that can be performed within our model framework. We are currently working on an extension that incorporates a budget constraint, as defined in Equation 6, rather than a total shock constraint. Such a constraint would allow for the inclusion of relevant costs incurred from strategic policy aimed at reducing certain cause-specific mortality rates. This would also allow us to consider socioeconomic-dependent shocks. Indeed, as some treatment or prevention plan may affect the groups differently, we may assume that public policy may wish to adjust their spending across specific groups in the population in order to achieve uniform shocks (or equity in outcomes) or, to the contrary, target resources on high-risk groups. In light of these costs, the optimal strategy might be adjusted in favor of others that maximize life expectancy "returns." In other words, the targeted cause of death would be identified not only via its impact on life expectancy and social inequality gaps, but also on the impact implementation would have on the national budget. Such a model would be of great added value for ensuring well-informed public policy.

Acknowledgment

Andrés M. Villegas acknowledges support by the Australian Research Council Centre of Excellence in Population Ageing Research (project number CE110001029). In addition, Madhavi Bajekal is funded by the longevity research collaboration between UCL and Legal & General Assurance Society Ltd (L&G).

References

Alai, D. H., Landsman, Z., and Sherris, M. (2013). Lifetime dependence modelling using a multivariate gamma distribution. *Insurance: Mathematics and Economics*, **52**(3), 542–549.

Alai, D. H., Arnold (-Gaille), S., and Sherris, M. (2015a). Modelling cause-of-death mortality and the impact of cause-elimination. *Annals of Actuarial Science*, **9**(1), 167–186.

Alai, D. H., Landsman, Z., and Sherris, M. (2015b). A multivariate Tweedie lifetime model: Censoring and truncation. *Insurance: Mathematics and Economics*, **64**, 203–213.

Alai, D. H., Landsman, Z., and Sherris, M. (2016a). Modelling lifetime dependence for older ages using a multivariate Pareto distribution. *Insurance: Mathematics and Economics*, **70**, 272–285.

Alai, D. H., Landsman, Z., and Sherris, M. (2016b). Multivariate Tweedie lifetimes: The impact of dependence. *Scandinavian Actuarial Journal*, **2016**(8), 692–712.

Anderson, R. N. (1999). US decennial life tables, 1989–91: United States life tables eliminating certain causes of death. *DHHS Publication No. (PHS) 99-1150-4*, **1**(4).

Arnold, S. and Sherris, M. (2013). Forecasting Mortality Trends Allowing for Cause-of-Death Mortality Dependence. *North American Actuarial Journal*, **17**(4), 273–282.

Arnold, S. and Sherris, M. (2015). Causes-of-Death Mortality: What Do We Know on Their Dependence? *North American Actuarial Journal*, **19**(2), 116–128.

Arnold, S. and Sherris, M. (2016). International Cause-Specific Mortality Rates: New Insights From a Cointegration Analysis. *ASTIN Bulletin: Journal of the International Actuarial Association*, **46**(1), 9–38.

Bajekal, M., Scholes, S., Blane, D., Lu, J., and Raine, R. (2011). Trends in social inequalities in old age life expectancy: Decomposition by age and cause of death (England, 1982 to 2006). *International Mortality and Longevity Symposium*. `http://www.actuaries.org.uk/research-and-resources/documents/c01-trends-inequalities-life-expectancy-age-65-england-socio-econom`.

Bajekal, M., Scholes, S. O., Flaherty, M., Raine, R., Norman, P., and Capewell, S. (2013a). Implications of using a fixed imd quintile allocation for small areas in England from 1981 to 2007. *PLOS ONE*, **8**(3).

Bajekal, M., Scholes, S. O., Flaherty, M., Raine, R., Norman, P., and Capewell, S. (2013b). Unequal trends in coronary heart disease mortality by socioeconomic circumstances, England 1982–2006: Analytical study. *PLOS ONE*, **8**(3).

Bayo, F. (1968). Life tables, 1959–61: United States life tables by causes of death, 1959–61. *Public Health Service Publication No 1252*, **1**(6).

Bennett, J. E., Li, G., Foreman, K., Best, N., Kontis, V., Pearson, C., Hambly, P., and Ezzati, M. (2015). The Future of Life Expectancy and Life Expectancy Inequalities in England and Wales: Bayesian Spatiotemporal forecasting. *Lancet*, **386**, 163–170.

Borooah, V. K. (2002). *Logit and Probit: Ordered and Multinominal Models*. Thousand Oaks, CA: Sage.

Branke, J., Deb, K., Dierolf, H., and Osswald, M. (2004). Finding Knees in Multi-objective Optimization. *Parallel Problem Solving From Nature*, pages 722–731.

Brønnum-Hansen, H. and Baadsgaard, M. (2012). Widening Social Inequality in Life Expectancy in Denmark: A Register-Based Study on Social Composition and Mortality Trends for the Danish Population. *BMC Public Health*, **12**(994).

Carriere, J. F. (1994). Dependent decrement theory. *Transactions, Society of Actuaries*, **46**, 45–65.

Chiang, C. L. (1968). *Introduction to Stochastic Process in Biostatistics*. New York: John Wiley and Sons.

Chiang, C. L. (1984). *The Life Table and Its Applications*. Malabar: Robert E. Krieger.

Curtin, L. R. and Armstrong, R. J. (1988). US decennial life tables, 1979-81: United States life tables eliminating certain causes of death. *DHHS Publication No. (PHS) 88-1150-2*, **1**(2).

De Grey, A. D. N. J. (2006). Extrapolaholics Anonymous: Why Demographers' Rejections of a Huge Rise in Cohort Life Expectancy in This Century Are Overconfident. In *Annals of the New York Academy of Sciences*, volume 1067, pages 83–93.

Department of Health (2003). Tackling Health Inequalities: A Programme for Action. Report 32366, Department of Health.

Dimitrova, D. S., Haberman, S., and Kaishev, V. K. (2013). Dependent competing risks: Cause elimination and its impact on survival. *Insurance: Mathematics and Economics*, **53**(2), 464–477.

Elandt-Johnson, R. C. (1976). Conditional failure time distributions under competing risk theory with dependent failure times and proportional hazard rates. *Scandinavian Actuarial Journal*, **1976**(1), 37–51.

Girosi, F. and King, G. (2006). *Demographic Forecasting*. Cambridge: Cambridge University Press.

Graham, H. (2004). Tackling inequalities in health in England: Remedying health disadvantages, narrowing health gaps or reducing health gradients? *Journal of Social Policy*, **33**(1), 115–131.

Greville, T. N. E., Bayo, F., and Foster, R. S. (1975). Life tables, 1969–71: United States life tables by causes of death, 1969–71. *DHEW Publication No. (HRA) 75-1150*, **1**(5).

Hayflick, L. (2004). "Anti-Aging" Is an Oxymoron. *Journal of Gerontology: Biological Sciences*, **59A**(6), 573–578.

Hougaard, P. (1984). Life table methods for heterogeneous populations: Distributions describing the heterogeneity. *Biometrika*, **71**(1), 75–83.

Kaishev, V. K., Dimitrova, D. S., and Haberman, S. (2007). Modelling the joint distribution of competing risks survival times using copula functions. *Insurance: Mathematics and Economics*, **41**(3), 339–361.

Lo, S. and Wilke, R. A. (2010). A copula model for dependent competing risks. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, **59**(2), 359–376.

Lorenc, T., Petticrew, M., Welch, V., and Tugwell, P. (2013). What Types of Interventions Generate Inequalities? Evidence From Systematic Reviews. *Journal of Epidemiology and Community Health*, **67**(2), 190–193.

Lu, J. L. C., Wong, W., and Bajekal, M. (2014). Mortality improvement by socio-economic circumstances in England (1982 to 2006). *British Actuarial Journal*, **19**(01), 1–35.

Mackenbach, J., Kulhnov, I., Artnik, B., Bopp, M., Borrell, C., Clemens, T., Costa, G., Dibben, C., Kalediene, R., Lundberg, O., Martikainen, P., Menvielle, G., stergren, O., Prochorskas, R., Rodrguez-Sanz, M., Heine Strand, B., Looman, C. W. N., and de Gelder, R. (2016). Changes in mortality inequalities over two decades: Register based study of European countries. *British Medical Journal*, **353**, i1732.

Mackenbach, J.-P., Kunst, A. E., Lautenbach, H., Oei, Y. B., and Bijlsma, F. (1999). Gains in Life Expectancy After Elimination of Major Causes of Death: Revised Estimates Taking Into Account the Effect of Competing Causes. *Journal of Epidemiology & Community Health*, **53**(1), 32–37.

Manton, K. G. (1986). Past and future life expectancy increases at later ages: Their implications for the linkage of morbidity, disability, and mortality. *Journal of Gerontology*, **41**(5), 672–681.

Manton, K. G. and Myers, G. C. (1987). Recent trends in multiple-caused mortality 1968 to 1982: Age and cohort components. *Population Research and Policy Review*, **6**, 161–176.

Manton, K. G. and Poss, S. S. (1979). Effects of dependency among causes of death for cause elimination life table strategies. *Demography*, **16**(2), 313–327.

Manton, K. G., Tolley, H. D., and Poss, S. S. (1976). Life table techniques for multiple-cause mortality. *Demography*, **13**(4), 541–564.

Manton, K. G., Stallard, E., and Poss, S. S. (1980). Estimates of U.S. multiple cause life tables. *Demography*, **17**(1), 85–102.

Manton, K. G., Stallard, E., and Vaupel, J. W. (1986). Alternative models for the heterogeneity of mortality risks among the aged. *Journal of the American Statistical Association*, **81**(395), 635–644.

Manton, K. G., Stallard, E., and Trolley, H. D. (1991). Limits to human life expectancy: Evidence, prospects, and implications. *Population and Development Review*, **17**(4).

Marmot Review Team (2010). *Fair Society, Healthy Lives: Strategic Review of Health Inequalities in England Post-2010*. London: University College London.

Menard, S. (2002). *Applied Logistic Regression Analysis*. Thousand Oaks, CA: Sage.

Noble, M., Mclennan, D., Wilkinson, K., Whitworth, A., Exley, S., Barnes, H., Dibben, C., McLennan, D., *et al.* (2007). The English indices of deprivation 2007.

Norman, P., Simpson, L., and Sabater, A. (2008). Estimating with confidence and hindsight: New UK small area population estimates for 1991. *Population, Space and Place*, **14**(5), 449–472.

Office for National Statistics (2015a). Inequality in Healthy Life Expectancy at Birth by National Deciles of Area Deprivation: England, 2011–2013. Technical report, Office for National Statistics.

Office for National Statistics (2015b). Trend in Life Expectancy at Birth and at Age 65 by Socio-economic Position Based on the National Statistics Socio-economic Classification, England and Wales: 1982–1986 to 2007–2011. Technical report, Office for National Statistics.

Olshansky, S. J., Hayflick, L., and Carnes, B. A. (2002). Position Statement on Human Aging. *Journal of Gerontology: Biological Sciences*, **57A**(8), B292–B297.

ONS (2015). Mortality statistics: Metadata. *Document of the Office for National Statistics*.

Rooney, C. and Devis, T. (1995). Mortality trends by cause of death in England and Wales 1980–94: The impact of introducing automated cause coding and related changes in 1993. *Population Trends*, **86**, 29–35.

Rooney, C., Griffiths, C., and Cook, L. (2002). The implementation of ICD-10 for cause of death coding: Some preliminary results from the bridge coding study. *Health Statistics Quarterly*, **13**, 31–41.

Rooney, C. I. F. and Smith, S. K. (2000). Implementation of ICD-10 for mortality data in England and Wales From January 2001. *Health Statistics Quarterly*, **8**, 41–51.

Rosén, M. (2006). Forecasting life expectancy and mortality in Sweden: Some comments on methodological problems and potential approaches. Technical Report 4, Social Insurance Studies From the Swedish Social Insurance.

United Nations (2011). United Nations General Assembly: Political Declaration of the High-Level Meeting of the General Assembly on the Prevention and Control of Non-communicable Diseases. `http://www.un.org/ga/search/view_doc.asp?symbol=A/66/L.1`.

Vaupel, J. W. and Yashin, A. I. (1983). The deviant dynamics of death in heterogeneous populations. Technical Report RR-83-001, International Institute for Applied Systems Analysis (IIASA).

Villegas, A. M. (2015). *Mortality: Modelling, Socio-Economic Differences and Basis Risk*. Ph.D. thesis, Cass Business School.

Villegas, A. M. and Haberman, S. (2014). On the modeling and forecasting of socioeconomic mortality differentials: An application to deprivation and mortality in England. *North American Actuarial Journal*, **18**(1), 168–193.

World Health Organization (1999). Health 21: The Health for All Policy Framework for the WEHO European Region. Technical report, WHO Regional Office for Europe.

World Health Organization (2008). Closing the Gap in a Generation: Health Equity Through Action on the Social Determinants of Health. Technical report, WHO Commission on Social Determinants of Health.

Appendix A Common and Prevalent Causes of Death

Table 6: ICD Classification

ICD Chapter	**Common Conditions**
Neoplasms (cancer)	Lung (trachea, bronchus), colon (rectum, anus), breast, prostate
Circulatory	Ischemic heart disease (heart attack), cerebrovascular disease (stroke)
Respiratory	Influenza and pneumonia, chronic obstructive respiratory disease (bronchitis, emphysema, asthma)
Digestive	Peptic ulcer, cirrhosis and chronic liver disease, hernia
External	Nontransport accidents, intentional self-harm, motor vehicle accidents
Other	Diabetes mellitus, dementia, renal failure, Parkinson's, Alzheimer's disease