

Adriaan Kalwij, Rob Alessie and Marike Knoef

**The Association between Individual
Income and Remaining Life Expectancy
at the Age of 65 in the Netherlands**

The Association between Individual Income and Remaining Life Expectancy at the Age of 65 in the Netherlands

By Adriaan Kalwij^{a d}, Rob Alessie^{b d} and Marike Knoef^{c d},

a) Corresponding author. Utrecht University, Janskerkhof 12, 3512 BL Utrecht, the Netherlands, Tel.: +31-30-2539664, Fax: +31-30-2537373, Email: a.s.kalwij@uu.nl

b) University of Groningen and Tinbergen Institute

c) Tilburg University and CentERdata

d) Network for Studies on Pensions, Aging and Retirement (Netspar)

Note: We wish to thank Margherita Borella, Koen Caminada, Katherine Carman, Stefan Hochguertel, Arie Kapteyn, Pierre Carl Michaud, and the seminar participants at the Utrecht School of Economics, the Center for Research on Pensions and Welfare policies in Turin (CeRP's 10th anniversary conference), University of Padova, the Netspar workshop "Labour Force Participation and Well-being of the 50⁺ Population" (September, 2009), the European Society for Population Economics conference 2010 and at the Econometric Society World Congress 2010 for valuable comments and discussions. Financial support has been provided by Stichting Instituut GAK through Netspar.

Abstract

This paper quantifies the association between individual income and remaining life expectancy at the statutory retirement age in the Netherlands of 65. For this purpose, we estimate a mortality risk model using a large administrative dataset that covers the 1996–2007 period. Besides age and marital status, the model includes as covariates individual and spouse's income, as well as a random individual specific effect. It thus allows for dynamic selection based on both observed and unobserved characteristics. We find that, conditional on marital status, individual income is about equally strong and negatively associated with mortality risk for men and women and that spouse's income is only weakly associated with mortality risk for women. For both men and women, we quantify that remaining life expectancy at age 65 for low-income individuals is about two-and-a-half years less than that for high-income individuals.

JEL codes: C23, C24, I10, J00

Keywords: Panel data, Dynamic selection, Mortality risk, Income

1. Introduction

Previous literature has identified significant socioeconomic inequalities in mortality risk over many populations and time periods (e.g., Duleep, 1986, Marmot et al., 1991, Menchik, 1993, Kunst and Mackenbach, 1994, Hurd et al., 1999, Attanasio and Hoynes, 2000, Blakely et al., 2004, Huisman, et al., 2004, Kunst et al., 2004, Palme and Sandgren, 2008, Hupfeld, 2009, Sullivan and von Wachter, 2009). These socioeconomic inequalities in mortality, commonly termed differential mortality, give strong evidence for an inverse relationship between income and mortality risk. An important economic implication of this inverse relationship is that, compared to high-income individuals, low-income individuals' internal rate of return from a uniformly priced pension plan is, on average, lower because their lower life expectancy results in receiving pension benefits for a relatively shorter period (Menchik, 1993, Brown, 2000, Simonovits, 2006, Bonenkamp, 2007, Hári, 2007).¹ This disparity is of concern to policymakers in developed countries like the Netherlands where a very large part of retirement consumption is financed by pension income (Knoef et al., 2009). For example, the recently proposed pension reforms take explicitly into account this disparity by facilitating workers in low-income sectors to receive a relative generous pension at most two years before the (newly proposed) statutory retirement age of 67 (Stichting van de Arbeid, 2010).²

¹ A closely related issue is that public pension policy in the Netherlands aims at redistributing income from the financially better off to the financially worse off, a redistribution that may be adversely affected by differential mortality (Nelissen, 1999).

² Pension income (mainly) consists of a state and an occupational pension. In case of earlier retirement (at most two years), state pension will be adjusted in an actuarially fair way. However, in collective agreements on a sector or firm level this loss of income can be compensated via the occupational

To gain insights into the size of the abovementioned differences in life expectancy between low and high-income individuals, we empirically quantify the association between individual income and remaining life expectancy at 65, the statutory retirement age in the Netherlands.³ In doing so, we draw on Grossman's (1972) seminal work on the relationships between socioeconomic position, health outcomes, and life expectancy, which serves as the theoretical framework for much of the empirical work in health economics. Specifically, Grossman's model assumes that individuals maximise lifetime utility and are endowed with a stock of knowledge capital that affects consumption (e.g., through investments in education), and a stock of health that deteriorates with age and impacts the time spent on market activities. Individuals derive utility from both the health stock and consumption, and the efficiency of investment in the health stock assumedly increases with knowledge capital. These mechanisms ensure that the health stock deteriorates with age at a relatively slower rate for individuals with a higher socioeconomic position (e.g., level of education). When it falls below a threshold value, the individual is deceased.⁴

Hence, in Grossman's model, an individual's socioeconomic position is positively related to both lifetime income (consumption) and life expectancy, which

pension. This compensation is made possible because in low-income sectors people have a below average life expectancy and pension premiums are based on population life expectancy.

³ Previous papers have been concerned with issues such as the implications of differential mortality for the age–wealth profile (Atkinson, 1971, Shorrocks, 1975), the incentives for saving during the working life for retirement consumption (Bloom et al., 2003), and the role of macroeconomic conditions in increasing life expectancy (Prichett and Summers, 1996, Cutler et al., 2006, Van den Berg et al., 2006). However, such analyses are not an aim of this paper.

⁴ Grossman (2000) discusses this aspect and gives an overview of the theoretical and empirical literature.

implies an inverse association between lifetime income and mortality risk. In line with this implication, empirical studies to date have identified an inverse relationship between individuals' financial situation and mortality risk. In fact, estimates of this relationship for different target populations with respect to country, age range, and different lifetime income approximations indicate that the ratio of mortality risk for individuals in the lowest quartile of the income or wealth distribution to that of individuals in the highest quartile ranges from around two in Europe to three in the U.S.⁵ At the same time, there is ongoing debate in the literature from various disciplines on the causal interpretation of this relationship and the possible pathways through which socioeconomic position affects health status and mortality (e.g., Marmot et al., 1991, Macintyre, 1997, Smith, 1999, Lindahl, 2005, Snyder and Evans, 2006). In this paper, although we take the position that there is, or might be, an association between income and mortality risk, we do not claim a causal relationship.

The contribution of this paper to the empirical literature on the relationship between income and mortality risk is threefold. First, in contrast to previous studies for the Netherlands that have most often examined the relation between education and mortality risk (e.g., Kunst and Mackenbach, 1994, Kippersluis et al., 2009), we estimate the association between income and mortality risk for individuals aged 65 or older. Moreover, we quantify the association between income and remaining life expectancy at age 65 by making use of Monte Carlo simulations. We choose age 65 because that is the statutory retirement age in the Netherlands after which all individuals receive (i) a public retirement pension that is independent of earnings

⁵ See Duleep (1986) and Attanasio and Hoynes (2000) for the U.S., Martikainen et al. (2001) for Finland, Osler et al. (2002) for Denmark, Attanasio and Emmerson (2003) for the U.K., Blakely et al. (2004) for New Zealand, and Gaudecker and Scholz (2006) for Germany.

history⁶ and (ii), an occupational pension that depends on earnings history (see e.g., Nelissen, 1999). Because this retirement income consists primarily of pension income and therefore closely related to individual earnings history, it serves as a good proxy for the lifetime income measure suggested by the theoretical literature discussed above.

Second, in contrast to most previous studies that, depending on data availability, have used either individual income or the sum of individual and spouse's income (i.e. household income), we make a distinction between individual income and spouse's income. In Grossman's model, with its focus on individual choice, a natural analytical starting point for men is the association between individual income and mortality risk. In this model, a spouse may improve the efficiency of an individual's investment in the health stock and the higher the partner's socioeconomic position, the greater the improvement. Indeed, for most women from the cohorts in our analysis, mortality risk is perhaps more likely to be negatively related to spouse's income than their own income because many have left the labor force at the time of marriage or birth of a first child. This issue is related to a frequent claim in the literature that household financial resources matter for health status and mortality risk (e.g., Martikainen et al., 2001), which in turn implies not only an association between individual income and mortality risk but also between spouse's income and mortality risk. We therefore test whether individual mortality for both men and women is associated with a spouse's income as well as the individual's income. Limited empirical evidence is available on this issue and the exceptions being McDonough et al. (1999) for the U.S. and Torssander and Erikson (2009) for Sweden who report that

⁶ Eligibility and amount depend only on the years of recorded residency in the Netherlands between the ages of 15 and 65 (2% of the full public pension benefit for each year).

for women, but not for men, spouse's income is negatively associated with individual mortality risk.

Third, our empirical model controls for unobserved individual-specific characteristics (i.e., random effects). This is of importance as it is inherent in the analysis of mortality risk that with age the sample becomes more selective in terms of both observed and unobserved characteristics. As a result, failing to control for this "dynamic selection" may bias the results (Cameron and Heckman, 1998, Van den Berg, 2001). Yet most of the studies cited above use (pooled) cross-sectional data, hence do not control for dynamic selection. Duration data make it possible to control for dynamic selection by including random effects. Nevertheless, using duration data, Hupfeld's (2009) model does not include random effects and Van den Berg et al. (2006) mention that including random effects does not affect the estimated impact of economic conditions early in life on individual mortality risk. Our empirical analysis uses panel data that are representative of the 65+ population and the model includes random effects to take into account dynamic selection. Nonetheless, a model complication does arise for individuals who enter the panel after age 65. That is, given the dynamic selection process, random effects at age 65 imply a dependency between these random effects and the covariates at later ages. To take this complication into account, we include dynamic sample selection correction terms that control for this dependency at the age of entry for individuals who enter the sample after age 65.

The paper is organized as follows. Section 2 describes the data. Section 3 outlines the empirical model for analyzing mortality risk and explains the estimation procedure. Section 4 reports the analytical results, and section 5 summarizes the main findings and concludes the paper.

2. Data

The data are taken from the 1996–2007 Income Panel Study of the Netherlands (IPO, *Inkomens Panel Onderzoek*, CBS 2009a) and the 1997–2008 Causes of Death registry (DO, *Doodsoorzaken*, CBS 2009b), both gathered by Statistics Netherlands. The IPO, a representative sample of the Dutch population, consists of an administrative panel dataset of about 92,000 individuals, randomly selected in 1996, which increased to about 99,000 individuals in 2007 because of population growth. Sampling is based on individuals' national security number, and the selected individuals are followed for as long as they are residing in the Netherlands on December 31 of the sample year. The dataset also includes individuals living in institutions for the elderly, such as nursing homes. Individuals born in the Netherlands enter the panel for the first time in the year of their birth; immigrants to the Netherlands, in the year of their arrival. An individual exits the panel on death or emigration from the Netherlands. Hence, the only reasons for panel attrition are mortality and emigration.⁷

The IPO contains data on the demographic characteristics and income of each member of a selected individual's household obtained from official institutions; most particularly, the population registry and tax office. The DO, on the other hand, provides information on date and cause of death for all residents deceased during the 1997–2008 period. These data come from medical records provided by medical

⁷ The annual rate of attrition in survey data tends to be 20% or more (e.g., Attanasio and Emmerson, 2003), meaning that a strong relationship between attrition and, for instance, health status may lead to inconsistent estimates of the association between income and mortality risk. A further drawback of survey data is that individuals living in institutions for the elderly are often not included.

examiners, who are legally obliged to submit them to Statistics Netherlands. The DO dataset also assigns a personal identifier that allows determination of whether an individual in the IPO has died by the next calendar year.

We select individuals aged 65 or over, who, because of population ageing, make up 12.8% of the sample in 1996 and 13.9% in 2007. This raw dataset consists of 151,120 observations for 21,159 individuals over the 1996–2007 period. We remove about 6% of the observations because of negative or zero income or missing values on the key variables (age, marital status and income), which affects relatively more men than women (11% vs. 4%). Nonetheless, once we control for gender, the mortality rate among the individuals excluded is not significantly different from the mortality rate among those included.⁸ Panel attrition for reasons other than mortality (i.e., because of emigration or missing values) is about 0.3% per year. The resulting sample consists of 19,258 individuals, 11,601 female and 7,657 male. Of these individuals, 42% enter the panel at the age of 65, while the remaining 58% enter the sample at a later age. The total sample contains 141,725 observations.

2.1 Variable Definitions and Descriptive Statistics

The analysis is based on the variables of gender, age, marital status, and income. We define age as the individual's age on January 1 of each year because in the Netherlands, the calendar year is also the fiscal year for income measurement, meaning that this choice ensures that income at age 65 is measured over the first entire calendar year of retirement. Table 1 reports the number of observations by age and gender and the distribution of marital status by age and gender. The marital status variable distinguishes between a single adult household that includes divorcees

⁸ The p -value corresponding to the null hypothesis of no difference is equal to 0.261.

(hereafter, “single”), a married or cohabiting couple (“married”), and a widowed individual. The differences in marital status across age and gender result from the recognized fact that women live on average longer than men. These differences result in, for instance, an increasing proportion of women with age (last column, panel A) and, at a given age, relatively more widowed and fewer married women than men (panel B).

The IPO income data are based primarily on tax records and contain detailed and accurate information on all income components. Here, income is gross of income tax and social insurance contributions and is measured in 2005 euros using the consumer price index. Individual income is the sum of pension, labor, transfer, and capital income.⁹ Table A1 (appendix A) provides an overview of these components and their definitions for both men and women and shows that over 90% of income is pension income. All income components are observed for the individual and, in couple households, also for the spouse. The analysis excludes any income from other household members.¹⁰

As shown in table 2,¹¹ the mean income of single and widowed men is higher than that of single and widowed women, and the distribution of income for single and widowed men is wider than the distribution of income for single and widowed women. The income distribution for men shows a decrease in median income with age. This finding most likely results from changes in income distribution over birth

⁹ The main results of this paper are unaffected by using a standardized (equivalized) income concept or using pension income only (a public retirement pension plus an occupational pension).

¹⁰ About 8% of households have other household members, mostly children. Excluding these households would not affect the main results of this paper.

¹¹ Following Statistics Netherlands guidelines, we do not report statistics based on a number of observations below 25. These statistics are designated in the tables by “-”.

cohorts (Knoef et al., 2009). In addition, the income distribution for married women is rather compressed, partly because many retired married women have no earnings history¹² and receive only public pension benefits but also because pre-1990, part-time work often came with no pension plan or a pension plan that had a relatively high threshold before contributions could be made.

A comparison of the tables for married men and women reveals that women's income accounts for, on average, one-third of household income. The rank correlation between the individual's and the spouse's income is about 0.10 (not shown in a table).

2.2 Differential Mortality

Defining mortality as being deceased in the following year, about 38% of the individuals died over the sample period. The results in table 3 confirm the accepted patterns that mortality risk increases with age and that men have a higher mortality risk than women. Likewise, age-specific mortality risk is lower among married individuals than among single or widowed individuals. In addition, the sample statistics on mortality risk by gender and age compare favorably with the population statistics reported in the HMD column; that is, the differences between the two by age and gender are small.

In table 4, panel A, we pool the results for single and widowed individuals as mortality risk by income quartile differs little between these two groups. Panel A shows that mortality risk decreases as income increases for single/widowed men and women, an effect that is strongest up to the third quartile. One measure of differential mortality is the ratio of mortality risk among individuals in the first quartile of the

¹² In our sample, about 25% of married women aged 65–69 receive an occupational pension; for women aged 70–74, this figure is about 20%, and for women aged 75 or over, it is about 16%.

income distribution to mortality risk among individuals in the fourth quartile (columns Q1/Q4). A comparison of panels A and B reveals that differential mortality is stronger for married men than for single/widowed men (2.4 vs. 1.7) but weaker for married women than for single/widowed women (1.4 vs. 1.9). These statistics are in line with findings of other European studies but lower than those reported for the U.S. (see the Introduction). Overall, differential mortality appears to decline with age. Interestingly however, as illustrated in panel C, when spouse's income is considered, a differential mortality pattern emerges for both men and women, one that, although weaker for men, is similar to that produced for individual income. We are aware of no comparable statistics in the literature.

3. Mortality Risk Model

This section outlines our empirical model for analyzing mortality risk. As discussed in section 2, we observe whether or not an individual is deceased in the subsequent year. Specifically, we consider the following latent variable model that relates next year's mortality risk, at age $(a+1)$, to individual's characteristics at age a :

$$H_{a+1} = -\alpha_a - X_a\beta - \Lambda - \varepsilon_a, \quad (1)$$

$$\begin{cases} M_{a+1} = 1 & \text{if } H_{a+1} < 0 \\ M_{a+1} = 0 & \text{otherwise} \end{cases}.$$

For now we suppress the index for the individual. In the context of this paper, and in line with Grossman's model, the latent variable H_{a+1} can be thought of as an individual's stock of health. If next year this stock falls below a certain threshold, normalized to zero in equation (1), the individual is deceased. The variable M_{a+1}

denotes observed mortality at age $(a+1)$, and M_{a+1} is equal to one if an individual became a years old and died at age $(a+1)$, and zero otherwise. α_a is an age-specific intercept, while X_a is a $(1 \times k)$ vector of an individual's observed characteristics at age a , including marital status and income, with a corresponding $(k \times 1)$ parameter vector β . Λ denotes an individual's unobserved characteristics, i.e. a random effect, and is assumed to be constant over time and independent of the covariates at age 65. This assumption does not, however, exclude dependency between the covariates and Λ at later ages. We also assume that the random effect is normally distributed with a zero mean and σ_Λ^2 variance and that the error term ε_a follows a logistic distribution and is independently distributed across individuals and time with a zero mean and a variance normalized to $\pi^2 / 3$.

3.1 Dynamic Selection

Inherent in any study of mortality risk over the lifecycle is recognition that the population at risk changes with age. By explicitly accounting for random effects, our model allows for sample selection by age on the basis of both observed and unobserved characteristics. Not accounting for such dynamic selection may yield inconsistent estimates of α_a and β (Cameron and Heckman, 1998, Van den Berg, 2001). If mortality risk is negatively related to income and positively to unobserved characteristic(s) (Λ), low-income (high-income) individuals with a low (high) Λ value are more likely to survive another year than low-income individuals with a high Λ value. Hence, in this example, dynamic selection results in a population at risk in which the correlation between Λ and income becomes increasingly positive with age up to a certain age and then decreases thereafter as the sample becomes more homogenous with respect to income and Λ .

When all individuals are observed from the age of 65, the model takes dynamic selection into account. However, as reported in section 2, 58% of the individuals enter the sample after age 65 and dynamic selection implies that these individuals, having survived from age 65 to the age of panel entry (τ), are a selective sample (of their cohort) in terms of their covariates (including income) and Λ . A post-65 entry thus produces the empirical complication of a dependency between the random effect and the covariates, a dependency that, as already explained, changes with the age of entry. Ideally one would like to control for this by explicitly modelling the probability of survival up to the age of entry (see, e.g., Ridder, 1984) but this would require data on the covariates from the age of 65 up to the age one year before entry, which we do not have. The solution we propose here is to explicitly account for the change in the dependency between random effects and income at the age of entry, thereby maintaining the random effects assumption at age 65. Essentially, using the formula below, we parametrize the change in the dependency between the covariates and the random effect Λ at the age of first observation τ for individuals who enter the sample after the statutory retirement age of 65:

$$\Lambda = \tilde{X}_\tau \gamma + \theta, \quad (2)$$

where $\tilde{X}_\tau = ((1, X_\tau) \times (\tau - 65), (1, X_\tau) \times (\tau - 65)^2)$ with a corresponding $((2k+2) \times 1)$ parameter vector γ . Scaling the effects of the covariates at age τ with the factors $(\tau - 65)$ and $(\tau - 65)^2$ takes into account, for instance, that the dependency between the random effect and income becomes more positive with age and at some age decreases as the sample becomes more homogeneous. θ is a random effect that is assumed to be independent of \tilde{X}_τ and normally distributed with a zero mean and σ^2 variance. Consistent estimates of the α_a 's and β are obtained under the additional

assumption formalized in equation (2). Testing the hypothesis $\sigma^2=0$ tests for the presence of random effects, while testing the hypothesis $\gamma=0$ tests for the need to include dynamic sample selection correction terms (i.e., the additional covariates \tilde{X}_τ).

3.2 Estimation, Identification, and Empirical Specification

Given the model outlined above, age-specific mortality risk conditional on observed and unobserved characteristics can be written as follows:

$$\Pr(M_{a+1} = 1 | X_\tau, X_a, M_a = 0) = \Pr(\varepsilon_a < \alpha_a + X_a\beta + \tilde{X}_\tau\gamma + \theta). \quad (3)$$

$$\Pr(M_{a+1} = 1 | X_\tau, X_a, M_a = 0) = F(\alpha_a + X_a\beta + \tilde{X}_\tau\gamma + \theta). \quad (4)$$

The condition $M_a = 0$ formalizes the fact that all individuals in the population at risk are alive at age a , and $F(\cdot)$ is the logistic cumulative distribution function. Equation (4) is the basis for a likelihood function in which we integrate out the unobserved individual specific effect. With i denoting the individual, $\tau(i)$ is the age of the individual when first observed in the sample and $A(i)$ is the age of the individual when last observed in the sample. The variable $m(i)$ is equal to one if the individual is deceased by age $A(i)+1$, and zero otherwise. Maximum likelihood estimates of the model parameters are given by

$$\begin{aligned} & (\hat{\alpha}, \hat{\beta}, \hat{\gamma}, \hat{\sigma}) = \\ & \operatorname{argmax}_{\alpha, \beta, \gamma, \sigma} \sum_{i=1}^N \log \left(\int_{-\infty}^{+\infty} \left(\prod_{a=\tau(i)}^{A(i)-1} \left(1 - F(\alpha_a + X_a(i)\beta + \tilde{X}_{\tau(i)}(i)\gamma + \theta(i)) \right) \right)^{I(A(i) > \tau(i))} \right. \\ & \quad \times \left(1 - F(\alpha_a + X_{A(i)}(i)\beta + \tilde{X}_{\tau(i)}(i)\gamma + \theta(i)) \right)^{1-m(i)} \\ & \quad \left. \times \left(F(\alpha_a + X_{A(i)}(i)\beta + \tilde{X}_{\tau(i)}(i)\gamma + \theta(i)) \right)^{m(i)} d\Phi \left(\frac{\theta(i)}{\sigma} \right) \right), \quad (5) \end{aligned}$$

where $\alpha = (\alpha_{65}, \dots, \alpha_T)^{13}$, N is the number of individuals, and Φ is the cumulative normal distribution function. The estimated model is often referred to as a random effects panel data logit model (Wooldridge, 2001). Equation (5) imposes proportionality between the age pattern, the covariates, and the random effect to ensure identification of the random effects distribution (see e.g., Cameron and Trivedi, 2005). For individuals who enter the sample at age 65 no dynamic sample selection correction terms are included (as for these $\tau=65$, hence $\tilde{X}_{\tau(i)}$ is a vector with zeros) and identification of the γ parameters is solely established by having individuals who enter the sample after age 65. We estimate the model separately for men and women.¹⁴

To estimate the association between the individual and spouse's income and mortality risk, we parametrize equation (1) as follows:

$$\begin{aligned}
 -H_{a+1} = & \alpha_0 + \alpha_1 \text{AGE}_a + \beta_1 \text{MARRIED}_a + \beta_2 \text{WIDOW}_a \\
 & + \beta_3 Y_a^I + \beta_4 Y_a^P + \beta_5 (Y_a^I)^2 + \beta_6 (Y_a^P)^2 + \beta_7 (Y_a^I Y_a^P) + \Lambda + \varepsilon_a.
 \end{aligned} \tag{6}$$

The associations between mortality risk and individual income (Y_a^I) and spouse's income (Y_a^P) are given by the parameters β_3 – β_7 , and spouse's income is equal to zero for a single or widowed individual. The main advantage of this specification is that it nests two empirical specifications used in the literature and discussed in the Introduction. The first nested model, which refers to the hypothesis $\beta_4 = \beta_6 = \beta_7 = 0$, is that in which only individual (and not spouse's) income has a direct association with mortality risk. The second nested model, which refers to the

¹³ T is the maximum age an individual may reach.

¹⁴ We use the *xtlogit* command of the software package STATA (www.stata.com) for estimation.

hypothesis $\beta_3 = \beta_4 \cap \beta_5 = \beta_6 \cap \beta_7 = 2\beta_5$, is that in which household income rather than individual and/or spouse's income is associated with mortality risk.

MARRIED is a dummy variable equal to one if the individual is married (including cohabitation) and zero otherwise; WIDOW is a dummy variable equal to one if the individual is widowed and zero otherwise. The reference category for marital status is a single adult household. Besides controlling for time effects, we also test for age-specific intercepts instead of a linear age function—which is briefly discussed in the next section.

3.3 Monte Carlo Simulations

Whereas the parameter estimates of the model outlined above provide insights into the direction and relative size of the associations between the covariates and mortality risk, they offer no clear insights into the quantitative association with remaining life expectancy at age 65. Therefore, the second part of the empirical analysis simulates remaining life expectancy at 65 and examines how it is associated with individual and spouse's income by gender and marital status at the age of 65. The model can thus be seen as a period-age model that conforms to the life tables used by Statistics Netherlands to calculate life expectancy based on the age-specific mortality risks of the current population (Van der Meulen and Janssen, 2007).¹⁵ We use the estimation results of the model outlined in sections 3.1 and 3.2 to simulate remaining life expectancy at age 65 (see appendix B for the technical details of the Monte Carlo simulations).

¹⁵ Theoretically, a cohort approach represents an alternative methodology. However, the available data do not allow for this method because, ideally, all individuals in the chosen cohorts should be followed up to the time of death.

Although income during retirement depends on lifetime earnings, the way that these two relate depends on the rules of the public and private pension systems. Hence, in the simulations, we take these rules into account when calculating income during retirement conditional on income before retirement (also known as pension-related gross yearly salary), reported in table 5 for several types of households. As explained earlier, the occupational pension income an individual receives in addition to the public pension is dependent on this gross pension-related salary. In the simulations we assume that individuals have lived in the Netherlands from the age of 15 onward and are therefore entitled to a flat rate public pension benefit that depends only on household composition. We also assume that, in the case of a married couple, the spouse's age is the same as the age of the individual. The baseline cases are individuals with a pension-related gross yearly salary equal to the median income in the Netherlands in 2005, which is €29,500 for full-time employees and €14,750 for part-time workers.¹⁶ We further assume that the occupational pension is based on 40 years of employment. In the cohorts included in the analysis, most men worked full time before the age of 65, but most women either did not work (or are not entitled to occupational pension benefits) or worked mostly part time (about a quarter) before the age of 65. Not only do the household types in table 5 take these situations into account, but even when other household types or different assumptions of spouse's age, employment history, and part-time earnings are considered, the main simulation results for deviations from the baseline are rather insensitive to (minor) changes in the baseline cases.

¹⁶ Source: <https://statline.cbs.nl> (Statistics Netherlands). We assume that a part-timer works half time with a median income equal to 50% of that of a full-time worker.

The rules we apply for calculating income during retirement conform to the rules applied by the largest pension funds (see the footnote to table 5). For instance, they take into account that when a woman becomes widowed, she is entitled to a part of the deceased husband's occupational pension. In the simulation exercise, for each household type, we consider differences from the baseline situation that result from changes in the pension-related gross yearly salary. Hence, in the simulation results, median income refers to median pension-related gross salary (see table 5), low income refers to a pension-related gross salary based on minimum wage throughout the working life, and high income refers to a pension-related gross salary based on twice the median income (see table A2, appendix A). These income classifications correspond roughly to the averages in the lowest and highest income quartiles for the different household types.

4. Empirical Results

The estimation results for the model outlined in section 3 are reported in table 6 and discussed in section 4.1. The simulation results are given in table 7 and discussed in section 4.2. In the following discussion we use a 5% level of significance. It is first worth noting, however, that we find no nonlinear age effects (in the index). We have examined this aspect using a model that contains a nonparametric specification of the age dependency of mortality risk and statistical tests have revealed that for both men

and women, the nonparametric age function can be restricted to a linear function of age.¹⁷

4.1 Estimation Results

An interpretation of the estimated associations in table 6 between marital status and mortality risk requires taking into account spouse's income, which is at least equal to the public pension benefit, and we return to this in section 4.2.

As shown in table 6, individual income is negatively associated with mortality risk for both men and women and the parameter estimates are roughly of the same magnitude. Based on these estimates, we calculate that the negative association between income and mortality risk diminishes after about €130,000 (which is around the top 0.2 percentile of the income distribution). Moreover, although the negative association between mortality risk and spouse's income is insignificant for men and only significant at a 10% level for women (see the second last row of table 6), it is of roughly the same magnitude for both. In addition, the test statistics for men and women indicate that we do not reject the hypothesis that it is household income rather than individual and/or spouse's income that is associated with mortality risk (see the last row of table 6).

The estimate of the standard deviation of the random effect is significant for both men and women, which suggests the presence of unobserved individual-specific heterogeneity (random effects). The standard deviations are, however, relatively small compared to the standard deviation of the error term in Eq. (1), which is equal to $\sqrt{\pi^2/3} \approx 1.81$. As shown in table 6, we also test for the importance of including

¹⁷ See appendix A. The p -values corresponding to an LR-test of model 2 against model 1 in table A3 and of model 6 against model 5 in table A4 are 0.33 and 0.18, respectively.

dynamic sample selection correction terms, an inclusion implied by the presence of random effects. These terms, as formulated in Eq. (2), are jointly significant for both men and women. Tables A3 (model 5) and A4 (model 10), appendix A, report the estimation results for the model that excludes random effects and dynamic sample selection correction terms. We do not discuss these in detail but conclude, in line with Van den Berg et al. (2006), that overall the differences with the estimates of table 6 are relatively small.

4.2 Simulation Results

The Monte Carlo simulations are based on the estimation results in table 6. These simulations, whose results are given in table 7, quantify the association between income and remaining life expectancy at the age of 65 in 2005 by gender for the different household types and income scenarios (see tables 5 and A2, appendix A). The point estimates of (unconditional) remaining life expectancy at age 65 are 16.0 years for men and 18.4 for women.¹⁸

As discussed in section 3, the baseline situation is for median income individuals and considers three types of households. For single men, remaining life expectancy is almost 12 years and for single women about 18 years (first column, first two rows of table 7). Remaining life expectancy at 65 is considerably higher for married individuals, about 16 years for married men (first column, third row) and 20 years for married women who have been employed part time (first column, last row).

¹⁸ For this purpose, we estimate the mortality risk model using only age and year dummies. The simulation-based estimates for remaining life expectancy are 15.99 (0.49) for men and 18.41 (0.38) for women.

The estimated differences between single and married individuals are significant: 4.4 years for men and 2.0 years for women.¹⁹

The second and third columns report on the change in remaining life expectancy associated with a 10% higher than median income for either the man or the woman in each household among the three household types. As table 7 shows, this association differs little across gender and household type: the point estimates vary between 0.16 and 0.20 with standard errors equal to 0.04. These estimates imply that life expectancy at age 65 for individuals with a 10% above median income is about two to two-and-a-half months higher than that for individuals with median income.

The fourth and fifth columns report the differences in remaining life expectancy between low and median income individuals for the different household types. An individual on minimum wage or with no earnings (low income) during the working life receives only a public retirement pension benefit during retirement. For both men and women, and irrespective of marital status at age 65, the difference in remaining life expectancy is less than one year (fourth and fifth columns). The differences in remaining life expectancy between high-income individuals, those who earned twice the median income during the working life, and median income individuals are reported in the sixth and seventh columns. For men and women, the point estimates of these differences fall between 1.53 (for married women) and 1.91 (for single men). A comparison of the differences between these two extremes reveals that the difference in remaining life expectancy at age 65 between low-income individuals with only a public retirement pension and high-income individuals with twice the median income is about two-and-a-half years for both men and women.

¹⁹ Estimates of these differences are 4.396 (0.515) for men and 2.032 (0.774) for women.

Finally, we turn to the association between remaining life expectancy at 65 and spouse's income. As shown in the fourth and sixth rows of table 7, compared to married women whose spouse earned median income, remaining life expectancy at age 65 of women whose spouse earned a 10% above the median income is 0.18-0.19 years higher (about two-and-a-half months). For married men (fifth line), however, this difference is small and insignificant. In contrast, the difference in remaining life expectancy at 65 between women with a low-income spouse and women with a high-income spouse (see columns four and six, last row) is more than two years ($1.62 - (-0.79)$). This difference is explainable in two ways. First, spouse's income is negatively associated with women's mortality risk, albeit only at a 10% level of significance (see table 6). Second, and more important, marriage is negatively associated with mortality risk and women benefit, on average, from an extended duration of marriage when married to a high-income man. For men, however, we find no such indirect association because, having on average a shorter life span than women, men benefit relatively less from a spouse's higher income.

5. Summary

This analysis quantifies the association between individuals' income and remaining life expectancy at the statutory retirement age in the Netherlands of 65. For this purpose, we estimate a mortality risk model that explicitly controls for unobserved individual-specific heterogeneity (random effects) using administrative data taken from the 1996–2007 Income Panel Study of the Netherlands supplemented with data from the Causes of Death registry.

Our main empirical findings are twofold. First, concerning model specification we find a significant presence of unobserved individual-specific heterogeneity (random effects) for both men and women. In our application the effects on the estimates when excluding random effects and the required dynamic sample selection correction terms are, however, relatively small.²⁰ The reason for this is that the estimated standard deviations of the random effect for men and women are relatively small and, in other words, random effects play only a minor role in our mortality risk model. Nevertheless, our findings underscore the importance of controlling for random effects and dynamic sample selection correction terms if inconsistent estimates are to be avoided.

Second, concerning the association of individual and spouse's income with mortality risk we find that, conditional on marital status, individual income is about equally strong and negatively associated with mortality risk for men and women and that spouse's income is only weakly (at a 10% level) associated with mortality risk for women. The difference in remaining life expectancy at age 65 between low-income individuals with only a public retirement pension and high-income individuals with twice the median income, is about two-and-a-half years for both men and women.

²⁰ For completeness, table A5 (appendix A) reports the simulation results based on a mortality model that does not control for random effects and dynamic sample selection. The differences with table 7 are also small.

References

- Atkinson, A.B., 1971, The distribution of wealth and the individual life-cycle, *Oxford Economic Papers*, New Series, Vol. 23, No. 2, pp. 239-254
- Attanasio, O.P., and Hoynes, H.W., 2000, Differential mortality and wealth accumulation, *Journal of Human Resources*, Vol. 35, No. 1, pp. 1-29
- Attanasio, O.P., and Emmerson, C., 2003, Mortality, health status and wealth, *Journal of the European Economic Association*, Vol. 1, No. 4, pp. 821-850
- Blakely, T., Kawachi, I., Atkinson, J, and Fawcett, J., 2004, Income and mortality: the shape of the association and confounding New Zealand Census-Mortality Study, 1981-1999, *International Journal of Epidemiology*, 33:8, 874-883
- Bonenkamp, J., 2007, Measuring lifetime redistribution in Dutch occupational pensions, *CPB discussion paper*, No. 82, Netherlands Bureau for Economic Policy Analysis.
- Brown, J.R., 2000, Differential mortality and the value of individual account retirement annuities, working paper 7560, National Bureau of Economic Research
- Cameron, S.V., and Heckman, J.J., 1998, Life cycle schooling and dynamic selection bias: models and evidence for five cohorts of American males, *Journal of Political Economy*, Vol. 106, No. 2, pp. 262-333.
- Cameron, A.C., and Trivedi, P.K., 2005. *Microeconometrics, Methods and Applications*. New York: Cambridge University Press
- Centraal Bureau voor de Statistiek (CBS), 2009a, Documentatierapport Inkomenspanel onderzoek (IPO), Centrum voor Beleidsstatistiek, Voorburg

- Centraal Bureau voor de Statistiek (CBS), 2009b, Documentatierapport Doodsoorzaken (DO), Centrum voor Beleidsstatistiek, Voorburg
- Cutler, D.M., Deaton, A.S., and Lleras-Muney, A., 2006, The determinants of mortality, Working paper 11963, National Bureau of Economic Research
- Duleep, H.O., 1986. Measuring income's effect on adult mortality using longitudinal administrative record data, *Journal of Human Resources*, 21(2), pp.238-251
- Gaudecker, von H.M., and Scholz, R.D., 2006, Lifetime earnings and life expectancy, *MPIDR working paper*, No. 2006-008, Max Planck Institute for Demographic Research
- Grossman, M., 1972. On the concept of health capital and the demand for health, *Journal of Political Economy*, Vol.80, No.2, pp. 223-255
- Grossman, M., 2000. The human capital model. In: Culyer, A.J., Newhouse, J.P. (Eds.), In: *Handbook of Health Economics*, vol. 1A. Elsevier Science, North-Holland, New York and Oxford, pp. 347–408.
- Hári, N., 2007, *Modeling mortality: Empirical studies on the effect of mortality on annuity markets*, Ph.D. Thesis, Tilburg University
- Huisman, M., Kunst, A.E., Andersen, O., Bopp, M., Borgan, J-K., Borrell, C., Costa, G., Deboosere, P., Desplanques, G., Donkin, A., Gadeyne, S., Minder, Ch., Regidor, E., Spadea, T., Valkonen, T., and Mackenbach, J.P., 2004, Socioeconomic inequalities in mortality among elderly people in 11 European populations, *Journal of Epidemiology and Community Health*, 58, pp. 468-475
- Human Mortality Database, 2008, <http://www.mortality.org>
- Hupfeld, S., 2009, Rich and healthy—better than poor and sick?: An empirical analysis of income, health, and the duration of the pension benefit spell, *Journal of Health Economics*, 28, pp. 427-443

- Hurd, M.D., McFadden, D., and Merrill, A., 1999, Predictors of mortality among elderly, *Working Paper 7440*, National Bureau of Economic Research
- Kippersluis, van H., O'Donell, O., and Van Doorslaer, E., 2009, Long run returns to education: does schooling lead to an extended old Age?, *Netspar Discussion paper* 04/2009-014
- Knoef, M., Alessie, R.J.M., and Kalwij, A., 2009, Changes in the income distribution of the Dutch elderly over the years 1989-2020: a microsimulation, *Netspar Discussion Paper* 09/2009 – 030 (www.netspar.nl)
- Kunst, A.E., and Mackenbach, J.P., 1994, The size of mortality differences associated with educational level in nine industrialized countries, *American Journal of Public Health*, 84, 932-937
- Kunst, A.E., Bos, V., Andersen, O., Cardano, M., Costa, G., g, S., Hemström, Ö., Layte, R., Regidor, E., Reid, A., Santana, P., Valkonen, T., and Mackenbach, J.P., 2004, Monitoring of trends in socioeconomic inequalities in mortality: Experiences from a European project, *Demographic Research*, Special Collection 2, pp. 229-254
- Law, A.M. and Kelton, W.D., 1982. *Simulation Modelling and Analysis*. New York: McGraw-Hill
- Lindahl, M., 2005, Estimating the effect of income on health and mortality using lottery prizes as an exogenous source of variation in income, *Journal of Human Resources*, Vol.40, No.1, pp. 144-168
- Macintyre, S., 1997, The Black Report and beyond what are the issues?, *Social Science & Medicine*, Vol. 44, No.6, pp. 723-745

- Marmot, M.G., Smith, G.D., Stansfeld, S., Patel, Ch., North, F., Head, J., White, I., Brunner, E., and Feeny, A., 1991, Health inequalities among British civil servants: The Whitehall II study, *Lancet*, June 8, pp. 1387-193
- Martikainen, P., Mäkelä, P., Koskinen, S., and Valkonen, T., 2001, Income differences in mortality: a register-based follow-up study of three million men and women, *International Journal of Epidemiology*, 30, pp. 1397-1405
- McDonough, P., Williams, D.R., and J.S. House, 1999, Gender and the socioeconomic gradient in mortality, *Journal of Health and Social Behavior*, Vol. 40, No. 1, pp. 17-31
- Menchik, P.L., 1993, Economic status as a determinant of mortality among black and white older men: Does poverty kill?, *Population Studies*, 47 pp. 427-436
- Nelissen, J.H.M., 1999, Mortality differences related to socioeconomic status and the progressivity of old-age pensions and health insurance: The Netherlands, *European Journal of Population*, 15, 77-97
- Osler, M., Prescott, E., Grønbaek, M., Christensen, U., Due, P., and Engholm, G., 2002. Income inequality, individual income, and mortality in Danish adults: analysis of pooled data from two cohort studies, *British Medical Journal*, Vol. 324, 13-16
- Palme, M., and Sandgren, S., 2008, Parental income, lifetime income, and mortality, *Journal of the European Economic Association*, 6(4), 890-911
- Prichett, L., and Summers, L.H, 1996, Wealthier is healthier, *The Journal of Human Resources*, Vol.31, No. 4, pp. 841-868
- Ridder, G., 1984, The distribution of single-spell duration data. In G.R. Neumann and N. Westergaard-Nielsen (eds.), *Studies in Labor Market Analysis*, Springer-Verlag, Berlin

- Shorrocks, A., 1975, The age-wealth Relationship: A cross-section and cohort analysis, *The Review of Economics and Statistics*, 57 (2), pp. 155-163
- Simonovits, A., 2006. Optimal design of pension rule with flexible retirement: The two-type case, *Journal of Economics*, Vol. 89, No.3, pp. 197-222
- Smith, J.P., 1999, Healthy bodies and thick wallets: The dual relation between health and economic status, *Journal of Economic Perspectives*, Vo. 13, No. 2, pp.145-166
- Snyder, S.E. and Evans, W.N., 2006, The effect of income on mortality: evidence from the social security notch, *The Review of Economics and Statistics*, 88 (3): 482-495
- Stichting van de Arbeid, 2010, *Pension Accord Spring 2010 (Pensioen Akkoord)*, www.stvda.nl
- Sullivan, D., and Von Wachter, T., 2009, Average earnings and long-term mortality: Evidence from administrative data, *American Economic Review: Papers & Proceedings*, 99:2, 133-138
- Torssander, J., and R. Erikson, 2009, Marital partner and mortality: the effects of social position of both spouses, *Journal of Epidemiology and Community Health*, 63, pp. 992-998
- Van den Berg, G.J. (2001) Duration models: Specification, identification, and multiple durations In: Heckman, JJ, Leamer, E (Eds), *Handbook of Econometrics*, Vol V North-Holland, Amsterdam
- Van den Berg, G., Lindeboom, M., and Portrait, F., 2006, Economic conditions early in life and individual mortality, *The American Economic Review*, 96, pp. 290-302

Van der Meulen, A., and Janssen, F., 2007, Achtergronden en berekeningswijzen van
CBS-overlevingstafels. Bevolkingstrends, 3^e kwartaal

Wooldridge, J.M., 2001, *Econometric Analysis of Cross Section and Panel Data*, The
MIT Press, Cambridge Massachusetts

Appendix A Tables

Table A1 Components of individual income during retirement ^{a)}

Male income components						
Age	Public pension %	Occupational pension %	Labor income %	Transfer income %	Capital income %	Total %
65-69	51.8	39.7	3.9	1.4	3.1	100.0
70-74	52.3	37.9	2.4	2.2	5.3	100.0
75-79	53.9	35.1	1.3	2.9	6.9	100.0
80-84	56.5	31.1	0.9	3.6	8.0	100.0
85-89	61.9	26.3	0.6	3.3	7.8	100.0
90-94	66.6	22.7	0.5	2.4	7.8	100.0
95+	67.5	25.8	1.6	0.9	4.1	100.0
Total	53.6	36.4	2.4	2.3	5.4	100.0

Female income components						
Age	Public pension %	Occupational pension %	Labor income %	Transfer income %	Capital income %	Total %
65-69	80.9	13.2	1.3	2.2	2.4	100.0
70-74	79.0	14.0	0.8	2.6	3.6	100.0
75-79	76.3	15.3	0.6	3.1	4.8	100.0
80-84	74.4	16.5	0.6	3.4	5.2	100.0
85-89	73.9	16.7	0.6	3.5	5.4	100.0
90-94	75.3	15.2	0.5	3.0	6.1	100.0
95+	75.7	15.4	0.2	3.0	5.7	100.0
Total	77.7	14.7	0.8	2.8	4.0	100.0

^{a)} All residents from the statutory retirement age of 65 onward receive a public retirement pension. Occupational pensions, on the other hand, are related to the individual's own employment or the employment of the spouse (in case of widowhood) before age 65. They therefore include private annuities, which may be more common among the self-employed. Labor income includes both work-related earnings and income from self-employment. Transfer income includes mainly alimony payments or receipts and rental subsidies. Capital income includes primarily interest, dividends, and income from real estate. A few income components might be considered household rather than individual income, e.g. rental subsidies. These components are recorded on an individual level by Statistics Netherlands. More details on the components of income and their definitions can be found in Knoef et al. (2009).

Table A2 Income scenarios during retirement for different household types

	Baseline situation		Difference from the baseline situation due to income differences				
	Man/Woman Median income	Man Median+10%	Woman Median+10%	Man Low income, Minimum wage	Woman Low income, Minimum wage	Man High income, 2x the median	Woman High income, 2x the median
	€	€	€	€	€	€	€
Pension-related gross yearly salary							
Full time	29500	32450		16392		59000	
Part time	14750		16225		8196		29500
Annual income during retirement by household type							
Single person household							
Man, was employed full time before age 65	20650	22715	22715	11705	11705	41300	41300
Woman, was employed part time before age 65	16178		17210		11705		26503
Couple, before age 65, the man was employed full time and the woman was not employed							
Man married at age 65	16962	19027		8018		37612	
Woman married at age 65	8018	8018		8018		8018	
Household income while married	24980	27045		16035		45630	
Man becomes widowed	20650	22715		11705		41300	
Woman becomes widowed	18094	19569		11705		32844	
Couple, before age 65, the man was employed full time and the woman part time							
Man married at age 65	16962	19027	16962	8018	16962	37612	16962
Woman married at age 65	12490	12490	13522	12490	8018	12490	22815
Household income while married	29452	31517	30485	20507	24980	50102	39777
Man becomes widowed	23845	25910	24582	14900	20650	44495	31220
Woman becomes widowed	22567	24042	23599	16178	18094	37317	32892

Table A3 Estimation results for men

Mortality risk	Model 1		Model 2		Model 3		Model 4		Model 5	
	Men		Men		Men		Men		Men	
	Parameter estimate	Standard error								
Constant	-3.336	0.206	-11.636	1.054	-11.398	1.037	-11.533	1.038	-10.468	0.240
Time dummy variables	included									
Age specific dummy variables ^{a)}	included		excluded		excluded		excluded		excluded	
Age (in years)			0.132	0.016	0.128	0.015	0.130	0.015	0.112	0.003
Single	0.000		0.000		0.000		0.000		0.000	
Married	-0.923	0.142	-0.872	0.142	-0.893	0.093	-0.753	0.096	-0.407	0.099
Widowed	-0.472	0.108	-0.436	0.104	-0.425	0.103	-0.428	0.104	-0.221	0.069
Y ¹ /10,000	-0.218	0.042	-0.214	0.041	-0.203	0.038			-0.190	0.024
Y ^P /10,000	-0.083	0.127	-0.097	0.124					-0.267	0.098
(Y ¹ /10,000) ²	0.008	0.002	0.008	0.002	0.009	0.002			0.009	0.001
(Y ^P /10,000) ²	-0.004	0.016	-0.003	0.015					0.017	0.008
(Y ¹ /10,000) x (Y ^P /10,000)	0.028	0.016	0.028	0.016					0.020	0.012
(τ -65)	-0.034	0.036	-0.053	0.034	-0.058	0.033	-0.048	0.033		
Single x (τ -65)	0.000		0.000		0.000		0.000			
Married x (τ -65)	0.167	0.046	0.154	0.044	0.090	0.025	0.098	0.026		
Widowed x (τ -65)	0.071	0.029	0.063	0.028	0.058	0.027	0.061	0.028		
(Y ¹ /10,000) x (τ -65)	-0.162	0.115	-0.159	0.111	-0.096	0.098				
(Y ^P /10,000) x (τ -65)	-1.025	0.543	-0.955	0.522						
(Y ¹ /10,000) ² x (τ -65)	0.017	0.009	0.017	0.009	0.017	0.008				
(Y ^P /10,000) ² x (τ -65)	0.098	0.099	0.094	0.094						
(Y ¹ /10,000) x (Y ^P /10,000) x (τ -65)	0.064	0.092	0.065	0.088						

Table A3: continued

	Model 1	Model 2	Model 3	Model 4	Model 5			
$(\tau-65)^2$	0.043	0.164	0.117	0.151	0.098	0.141	0.063	0.143
Single x $(\tau-65)^2$	0.000		0.000		0.000		0.000	
Married x $(\tau-65)^2$	-0.725	0.258	-0.659	0.246	-0.364	0.126	-0.416	0.131
Widowed x $(\tau-65)^2$	-0.291	0.146	-0.251	0.137	-0.230	0.136	-0.241	0.137
$(Y^l/10,000) \times (\tau-65)^2$	0.075	0.063	0.078	0.060	0.084	0.053		
$(Y^p/10,000) \times (\tau-65)^2$	0.373	0.338	0.347	0.323				
$(Y^l/10,000)^2 \times (\tau-65)^2$	-0.009	0.006	-0.009	0.006	-0.011	0.005		
$(Y^p/10,000)^2 \times (\tau-65)^2$	-0.037	0.053	-0.035	0.051				
$(Y^l/10,000) \times (Y^p/10,000) \times (\tau-65)^2$	0.023	0.060	0.018	0.057				
$Y^H/10,000$							-0.196	0.038
$(Y^H/10,000)^2$							0.008	0.002
$(Y^H/10,000) \times (\tau-65)$							-0.165	0.103
$(Y^H/10,000)^2 \times (\tau-65)$							0.017	0.008
$(Y^H/10,000) \times (\tau-65)^2$							0.102	0.056
$(Y^H/10,000)^2 \times (\tau-65)^2$							-0.009	0.005
Standard deviation, random effect	0.523	0.365		0.340			0.346	
Log-likelihood value	-10872.5	-10888.9		-10902.6			-10897.1	-10914.8
Number of observations	54617	54617		54617			54617	54617
Number of individuals	7657	7657		7657			7657	7657
Number of parameters	66	36		27			27	19

a) Dummy variables are included for age 66 up to and including age 95 and for a linear effect for ages over 95 (with 65 as the reference age). When these are excluded, the age variable is included.

Table A4 Estimation results for women

Mortality risk	Model 6	Model 7	Model 8	Model 9	Model 10					
	Women	Women	Women	Women	Women					
	<i>Parameter estimate</i>	<i>Standard error</i>								
Constant	-4.197	0.220	-11.337	0.891	-11.501	0.886	-11.434	0.888	-12.155	0.215
Time dummy variables	included									
Age specific dummy variables ^{a)}	included		excluded		excluded		excluded		excluded	
Age (in years)			0.109	0.014	0.112	0.014	0.109	0.014	0.121	0.002
Single	0.000		0.000		0.000		0.000		0.000	
Married	-0.359	0.140	-0.352	0.136	-0.529	0.110	-0.153	0.110	-0.301	0.090
Widowed	0.018	0.111	0.014	0.107	0.005	0.106	0.006	0.106	-0.037	0.050
Y ¹ /10,000	-0.282	0.050	-0.271	0.048	-0.270	0.045			-0.186	0.030
Y ^P /10,000	-0.104	0.057	-0.103	0.056					-0.091	0.040
(Y ¹ /10,000) ²	0.011	0.003	0.010	0.003	0.011	0.003			0.008	0.002
(Y ^P /10,000) ²	0.002	0.004	0.002	0.004					0.003	0.003
(Y ¹ /10,000) x (Y ^P /10,000)	0.016	0.022	0.015	0.022					-0.003	0.016
(τ-65)	0.016	0.033	0.020	0.030	0.019	0.029	0.030	0.028		
Single x (τ-65)	0.000		0.000		0.000		0.000			
Married x (τ-65)	-0.037	0.030	-0.029	0.028	-0.023	0.022	-0.037	0.022		
Widowed x (τ-65)	-0.025	0.021	-0.022	0.020	-0.020	0.020	-0.020	0.020		
(Y ¹ /10,000) x (τ-65)	0.086	0.121	0.094	0.114	0.076	0.109				
(Y ^P /10,000) x (τ-65)	0.207	0.163	0.198	0.155						
(Y ¹ /10,000) ² x (τ-65)	0.009	0.012	0.008	0.011	0.008	0.011				
(Y ^P /10,000) ² x (τ-65)	0.005	0.011	0.005	0.010						
(Y ¹ /10,000) x (Y ^P /10,000) x (τ-65)	-0.199	0.115	-0.187	0.109						

Table A4: continued

	Model 6	Model 7	Model 8	Model 9	Model 10			
$(\tau-65)^2$	-0.184	0.117	-0.084	0.103	-0.095	0.101	-0.094	0.092
Single x $(\tau-65)^2$	0.000		0.000		0.000		0.000	
Married x $(\tau-65)^2$	0.325	0.162	0.281	0.150	0.207	0.102	0.237	0.100
Widowed x $(\tau-65)^2$	0.125	0.086	0.106	0.080	0.102	0.080	0.101	0.079
$(Y^I/10,000) \times (\tau-65)^2$	-0.002	0.058	-0.007	0.054	0.003	0.053		
$(Y^P/10,000) \times (\tau-65)^2$	-0.184	0.115	-0.171	0.108				
$(Y^I/10,000)^2 \times (\tau-65)^2$	-0.007	0.006	-0.006	0.006	-0.006	0.006		
$(Y^P/10,000)^2 \times (\tau-65)^2$	0.000	0.009	0.000	0.009				
$(Y^I/10,000) \times (Y^P/10,000) \times (\tau-65)^2$	0.145	0.082	0.134	0.076				
$Y^H/10,000$							-0.187	0.038
$(Y^H/10,000)^2$							0.006	0.002
$(Y^H/10,000) \times (\tau-65)$							0.042	0.082
$(Y^H/10,000)^2 \times (\tau-65)$							0.003	0.006
$(Y^H/10,000) \times (\tau-65)^2$							-0.006	0.040
$(Y^H/10,000)^2 \times (\tau-65)^2$							-0.002	0.003
Standard deviation, random effect	0.507	0.233		0.233			0.232	
Log-likelihood value	-14804.9	-14823.3		-14830.4			-14831.1	-14836.5
Number of observations	87108	87108		87108			87108	87108
Number of individuals	11601	11601		11601			11601	11601
Number of parameters	66	36		27			27	19

a) Dummy variables are included for age 66 up to and including age 95 and for a linear effect for ages over 95 (with 65 as the reference age). When these are excluded, the age variable is included.

Table A5 Simulation results when random effects and dynamic selection correction terms are excluded. Estimation results are given in tables A3 (model 5) and table A4 (model 10).

	Baseline situation	Difference from the baseline situation due to differences in lifetime income					
	Median income	Man Median+10%	Woman Median+10%	Man Low income	Woman Low income	Man High income	Woman High income
Type of household	<i>Remaining life expectancy</i>	<i>Difference in remaining life expectancy</i>	<i>Difference in remaining life expectancy</i>	<i>Difference in remaining life expectancy</i>	<i>Difference in remaining life expectancy</i>	<i>Difference in remaining life expectancy</i>	<i>Difference in remaining life expectancy</i>
<i>Single person household</i>							
Man	13.00 (0.52)	0.20 (0.03)		-0.90 (0.13)		1.86 (0.29)	
Woman	17.30 (0.44)		0.11 (0.02)		-0.50 (0.09)		1.09 (0.17)
<i>Couple, before age 65, the man was employed full time and the woman was not employed</i>							
Man	15.99 (0.44)	0.20 (0.03)		-0.92 (0.12)		1.83 (0.26)	
Woman	18.74 (0.38)	0.14 (0.03)		-0.64 (0.13)		1.26 (0.27)	
<i>Couple, before age 65, the man was employed full time and the woman part time</i>							
Man	16.51 (0.46)	0.19 (0.03)	0.11 (0.04)	-0.88 (0.12)	-0.53 (0.18)	1.74 (0.25)	1.09 (0.41)
Woman	19.25 (0.36)	0.13 (0.03)	0.11 (0.02)	-0.62 (0.13)	-0.51 (0.08)	1.24 (0.27)	1.09 (0.18)

* Tables 5 and A2 provide details on the income classifications. Standard errors are in parentheses.

Appendix B Monte Carlo simulations

The Monte Carlo simulations are based on the distribution function of remaining life duration at age 65 (see Appendix C) and carried out as follows. First, as a baseline, we define a group of reference individuals with the same characteristics; for instance, married men with median income. The estimates of the parameters of Eq. (1) together with the assumption that the error term in Eq. (1) follows a logistic distribution enable calculation of the probabilities that at age 65 each individual in this reference group will be deceased by the subsequent year. Next, we compare these probabilities with random drawings from the uniform distribution to simulate whether or not each individual is deceased in the subsequent year (see e.g., Law and Kelton, 1982). Finally, we simulate age by age the mortality status for each individual in the reference group up to the age of 105. We assume that next year's mortality probability is equal to one at age 105. In this way, we obtain the simulated mortality status for each individual in this group from age 65 up to age 105. We use these sequences to calculate the mean remaining life duration at 65 for this homogenous reference group of individuals. We then rerun these simulations with a change in one of the covariates (e.g., a different income level) so that, all other factors being constant, the differences between these simulation outcomes and the baseline simulation outcomes can reveal this covariate's association with remaining life expectancy at 65. For reference individuals who are married at age 65 the simulations take explicitly into account that the spouse may die and the individual becomes widowed.

We perform these Monte Carlo simulations for 10,000 (identical) individuals. The standard errors for the differences from the baseline situation are based on 100 drawings from the asymptotic distribution of the parameter estimates.

Appendix C The distribution function of remaining life duration

The distribution function of remaining life duration at age τ , denoted by D , conditional on a vector of observed characteristics $X = \{X_\tau, \dots, X_d\}$ is modelled as a sequence of yearly mortality risks and for $d > \tau$ is given by

$$\begin{aligned}
 & \Pr(D = d \mid D \geq \tau; X, \theta) \\
 &= \frac{\Pr(D = d \cap D \geq \tau \mid x, \theta)}{\Pr(D \geq \tau \mid x, \theta)} = \frac{\Pr(D = d \mid x, \theta)}{\Pr(D \geq \tau \mid x, \theta)}, \\
 & \text{,,} \\
 &= \frac{\Pr(D \geq d \mid x, \theta)}{\Pr(D \geq \tau \mid x, \theta)} \times \frac{\Pr(D = d \mid x, \theta)}{\Pr(D \geq d \mid x, \theta)}, \\
 & \text{,,} \\
 &= \frac{\Pr(D \geq d \mid x, \theta)}{\Pr(D \geq \tau \mid x, \theta)} \times \Pr(D = d \mid D \geq d; X, \theta), \\
 & \text{,,} \\
 &= \frac{\Pr(D \geq d \mid x, \theta)}{\Pr(D \geq \tau \mid x, \theta)} \times F(\alpha_a + X_d \beta + \tilde{X}_\tau \gamma + \theta), \\
 & \text{,,} \\
 &= \frac{\prod_{a=0}^{d-1} (1 - F(\alpha_a + X_a \beta + \tilde{X}_\tau \gamma + \theta))}{\prod_{a=0}^{\tau-1} (1 - F(\alpha_a + X_a \beta + \tilde{X}_\tau \gamma + \theta))} \times F(\alpha_a + X_d \beta + \tilde{X}_\tau \gamma + \theta), \\
 & \text{,,} \\
 &= \left[\prod_{a=\tau}^{d-1} (1 - F(\alpha_a + X_a \beta + \tilde{X}_\tau \gamma + \theta)) \right] \times F(\alpha_a + X_d \beta + \tilde{X}_\tau \gamma + \theta),
 \end{aligned}$$

and for $d = \tau$ given by

$$\begin{aligned}
 & \Pr(D = d \mid D \geq \tau; X, \theta) \\
 &= \frac{\Pr(D = d \mid x, \theta)}{\Pr(D \geq d \mid x, \theta)} = \frac{\Pr(D \geq d \mid x, \theta)}{\Pr(D \geq d \mid x, \theta)} \times \Pr(D = d \mid D \geq d; X, \theta), \\
 & \text{,,} \\
 &= F(\alpha_a + X_d \beta + \tilde{X}_\tau \gamma + \theta).
 \end{aligned}$$

Table 1 Number of observations by age, and marital status by age and gender

Panel A	Number of observations		Age distribution by gender		
	Men	Women	Men	Women	Share of women
Age			%	%	%
65-69	17,504	23,349	32.0	26.8	57.2
70-74	15,943	21,558	29.2	24.7	57.5
75-79	11,530	18,593	21.1	21.3	61.7
80-84	6,305	13,307	11.5	15.3	67.9
85-89	2,577	7,085	4.7	8.1	73.3
90-94	664	2,692	1.2	3.1	80.2
95+	94	524	0.2	0.6	84.8
All	54,617	87,108	100.0	100.0	61.5

Panel B	Marital status				Marital status			
	Men				Women			
	Single	Widowed	Married	All	Single	Widowed	Married	All
Age	%	%	%	%	%	%	%	%
65-69	14.3	7.0	78.7	100.0	13.1	24.1	62.8	100.0
70-74	10.4	10.7	78.9	100.0	12.5	37.2	50.3	100.0
75-79	8.7	17.1	74.2	100.0	11.8	52.4	35.8	100.0
80-84	8.3	26.2	65.5	100.0	11.7	66.6	21.7	100.0
85-89	7.8	41.3	50.9	100.0	12.1	77.4	10.6	100.0
90-94	6.3	59.9	33.7	100.0	14.1	81.6	4.3	100.0
95+	4.3	62.8	33.0	100.0	16.0	83.0	1.0	100.0
All	10.9	14.8	74.4	100.0	12.4	46.4	41.2	100.0

Table 2 Distribution of individual income by marital status, age, and gender

	Single and widowed men				Single and widowed women			
	Mean	25th percentile	Median	75th percentile	Mean	25th percentile	Median	75th percentile
Age	€	€	€	€	€	€	€	€
65-69	22,283	14,454	17,942	25,180	19,894	13,827	16,270	21,757
70-74	22,143	14,628	18,130	24,925	19,684	13,926	16,374	21,380
75-79	22,945	14,474	18,198	26,077	19,407	13,829	16,145	20,621
80-84	22,342	13,710	17,079	24,921	19,115	13,524	15,560	19,905
85-89	20,387	13,151	15,957	22,885	18,674	13,069	15,046	19,215
90-94	20,797	12,944	15,341	23,556	18,428	12,411	14,586	18,351
95+	18,210	12,868	15,213	20,000	19,055	12,203	14,169	18,249
All	22,163	14,194	17,640	25,045	19,345	13,597	15,843	20,518
	Married men				Married women			
	Mean	25th percentile	Median	75th percentile	Mean	25th percentile	Median	75th percentile
Age	€	€	€	€	€	€	€	€
65-69	23,699	13,480	18,021	27,595	9,558	7,862	8,088	9,050
70-74	21,463	12,434	16,376	24,881	9,242	7,882	8,088	8,672
75-79	20,715	11,831	15,563	23,995	9,294	7,882	8,088	8,572
80-84	19,930	11,373	14,787	23,610	9,447	7,882	8,088	8,657
85-89	19,182	10,777	13,602	22,028	9,639	7,882	8,104	8,957
90-94	16,846	10,000	12,119	16,854	9,476	7,983	8,104	9,197
95+	22,604	9,984	12,702	32,184	-	-	-	-
All	21,809	12,391	16,469	25,324	9,407	7,882	8,089	8,818

Table 3 Mortality risk by gender, age, and marital status. The HMD columns provide population statistics from the Human Mortality Database (2008) over the 1996–2006 period.

	Mortality risk, men					Mortality risk, women				
	Single	Widowed	Married	All	HMD	Single	Widowed	Married	All	HMD
Age	%	%	%	%	%	%	%	%	%	%
65-69	4.6	3.4	1.5	2.1	2.2	1.6	1.7	1.0	1.3	1.2
70-74	5.8	5.5	3.5	3.9	3.8	2.7	2.5	1.6	2.1	2.0
75-79	8.3	8.6	6.1	6.7	6.4	4.9	4.3	3.0	3.9	3.5
80-84	13.6	13.0	9.8	11.0	10.8	8.2	8.1	6.8	7.9	6.6
85-89	23.8	18.5	15.9	17.6	17.5	12.8	13.6	10.8	13.2	12.4
90-94	33.3	30.2	24.1	28.3	27.5	20.3	20.3	19.8	20.3	21.7
95+	-	33.9	32.3	35.1	41.8	33.3	35.2	-	34.9	36.2
All	7.2	10.6	4.5	5.7	5.7	5.3	6.9	2.3	4.8	4.7

Table 4 Mortality risk by age, gender, and income quartile^{a)}

Cells: Mortality risk										
Panel A Single and widowed men						Single and widowed women				
	Income quartile					Income quartile				
	Q1	Q2	Q3	Q4	Q1/Q4	Q1	Q2	Q3	Q4	Q1/Q4
Age	%	%	%	%	%	%	%	%	%	%
65-69	5.5	4.2	3.4	3.8	1.4	2.7	1.2	1.4	1.4	1.9
70-74	6.6	7.5	5.2	3.5	1.9	3.8	2.5	2.4	1.8	2.1
75-79	12.0	8.2	6.9	7.3	1.6	6.2	4.7	3.5	3.5	1.8
80-84	16.4	12.1	12.2	11.4	1.4	10.1	7.7	7.1	7.5	1.3
85-89	22.1	19.1	18.8	15.6	1.4	17.8	11.3	11.2	12.1	1.5
90-94	35.7	30.8	27.0	23.5	1.5	22.3	22.4	17.9	16.3	1.4
95+	-	-	-	-	-	40.3	33.7	26.7	30.0	1.3
All	12.6	9.3	7.5	7.3	1.7	9.8	6.2	5.0	5.2	1.9

Panel B Married men						Married women				
	Income quartile					Income quartile				
	Q1	Q2	Q3	Q4	Q1/Q4	Q1	Q2	Q3	Q4	Q1/Q4
Age	%	%	%	%	%	%	%	%	%	%
65-69	3.2	1.8	1.0	1.0	3.3	1.5	1.0	0.8	0.7	2.1
70-74	4.6	3.6	3.3	2.3	2.0	2.3	1.2	1.5	1.3	1.7
75-79	7.2	6.5	5.4	4.6	1.6	3.4	2.9	2.5	3.2	1.1
80-84	11.3	10.0	8.5	8.2	1.4	7.6	7.2	6.9	5.3	1.4
85-89	17.1	13.4	16.5	15.5	1.1	11.8	10.1	9.5	11.8	1.0
90-94	28.3	20.4	21.4	14.7	1.9	34.8	14.8	6.3	26.5	1.3
95+	-	-	-	-	-	-	-	-	-	-
All	7.1	4.8	3.4	3.0	2.4	2.8	2.2	2.1	2.0	1.4

Panel C Married men						Married women				
	Spouse's income quartile					Spouse's income quartile				
	Q1	Q2	Q3	Q4	Q1/Q4	Q1	Q2	Q3	Q4	Q1/Q4
Age	%	%	%	%	%	%	%	%	%	%
65-69	1.8	1.3	1.6	1.4	1.3	1.3	1.1	1.1	0.8	1.7
70-74	4.3	3.1	3.1	3.3	1.3	2.1	1.1	1.7	1.6	1.3
75-79	7.2	6.4	5.1	5.5	1.3	2.9	3.1	3.6	2.7	1.1
80-84	11.8	9.2	8.9	9.6	1.2	8.2	7.5	5.5	5.5	1.5
85-89	16.2	17.4	15.5	14.5	1.1	13.2	6.1	11.0	11.0	1.2
90-94	25.4	32.5	15.6	26.3	1.0	27.5	16.7	19.4	10.7	2.6
95+	-	-	-	-	-	-	-	-	-	-
All	5.3	4.4	4.4	4.0	1.3	3.1	2.0	2.3	1.9	1.6

^{a)} The income quartiles are reported in table 2.

Table 5 Income during retirement for different household types

Cells: Yearly amounts in 2005	€
Statutory minimum wage, full time	16392
Public pension for a single person household	11705
Public pension per person for a two person household	8018
	Baseline situation
Pension-related gross yearly salary, median	€
Man (full time)	29500
Woman (part time)	14750
Annual income during retirement by household type^{a)}	€
<i>Single person household</i>	
Man (employed full time before age 65)	20650
Woman (employed part time before age 65)	16178
<i>Couple, before age 65, the man was employed full time and the woman was not employed</i>	
Man married at age 65	16962
Woman married at age 65	8018
Household income while married	24980
Man is widowed	20650
Woman is widowed	18094
<i>Couple, before age 65, the man was employed full time and the woman part time</i>	
Man married at age 65	16962
Woman married at age 65	12490
Household income while married	29452
Man is widowed	23845
Woman is widowed	22567

^{a)} Pension-related gross salary (Y) refers to the gross salary on which the occupational pension is based, that is received in addition to a public pension. Occupational pension income is calculated using the formula $\max(0, 0.7 \times \text{FTE} \times (Y - 10/7 \times \text{FRANCHISE}))$. A widowed individual receives $\max(0, 5/7 \times 0.7 \times \text{FTE} \times (Y_p - 10/7 \times \text{FRANCHISE}))$, where Y_p refers to the pension-related gross salary of the deceased spouse. FRANCHISE is set equal to €11705 (the public retirement pension received by a person living alone). FTE is equal to one if the individual was employed full-time and 0.5 if employed part-time.

Table 6 Estimation results*

Dependent variable: Mortality risk	Men		Women	
	<i>Parameter estimate</i>	<i>Standard error</i>	<i>Parameter estimate</i>	<i>Standard error</i>
<i>Covariate, parameter, Eq. (6)</i>				
Constant, α_0	-11.636	1.054	-11.337	0.891
Age, α_1	0.132	0.016	0.109	0.014
Single	0.000		0.000	
Married, β_1	-0.872	0.142	-0.352	0.136
Widowed, β_2	-0.436	0.104	0.014	0.107
(Individual income/10,000), β_3	-0.214	0.041	-0.271	0.048
(Spouse's income/10,000), β_4	-0.097	0.124	-0.103	0.056
(Individual income/10,000) ² , β_5	0.008	0.002	0.010	0.003
(Spouse's income/10,000) ² , β_6	-0.003	0.015	0.002	0.004
(Individual income/10,000) x (Spouse's income/10,000), β_7	0.028	0.016	0.015	0.022
Standard deviation random effect, σ	0.365	0.176	0.233	0.047
Log-likelihood value	-10888.9		-14823.3	
Number of observations	54617		87108	
Number of individuals	7657		11601	
Number of parameters	36		36	
<i>Test of hypothesis</i>	<i>p-value</i>		<i>p-value</i>	
Dynamic sample selection correction terms ^{a)}	0.0002		0.034	
Individual income ^{b)}	0.000		0.000	
Spouse's income ^{c)}	0.554		0.091	
Household income vs. individual/spouse's income ^{d)}	0.709		0.356	

* The estimates of all model parameters are reported in tables A3 (model 2) and A4 (model 7), appendix A.

^{a)} $H_0: \gamma=0$ (Eq. 2) ^{b)} $H_0: \beta_3 = \beta_5 = 0$ ^{c)} $H_0: \beta_4 = \beta_6 = \beta_7 = 0$ ^{d)} $H_0: \beta_3 = \beta_4, \beta_5 = \beta_6, \beta_7 = 2\beta_5$

Table 7 Simulations of remaining life expectancy at the statutory retirement age of 65 (in years) by gender and household type. These are based on the estimation results of table 6*

	Baseline situation	Difference from the baseline situation due to differences in lifetime income					
	Median income	Man Median+10%	Woman Median+10%	Man Low income	Woman Low income	Man High income	Woman High income
Type of household	Remaining life expectancy	Difference in remaining life expectancy	Difference in remaining life expectancy	Difference in remaining life expectancy	Difference in remaining life expectancy	Difference in remaining life expectancy	Difference in remaining life expectancy
<i>Single person household</i>							
Man	11.73 (0.66)	0.20 (0.04)		-0.88 (0.17)		1.91 (0.39)	
Woman	18.18 (1.14)		0.18 (0.04)		-0.80 (0.17)		1.74 (0.36)
<i>Couple, before age 65, the man was employed full time and the woman was not employed</i>							
Man	16.12 (0.57)	0.20 (0.04)		-0.91 (0.18)		1.86 (0.38)	
Woman	19.50 (0.96)	0.19 (0.06)		-0.82 (0.23)		1.71 (0.50)	
<i>Couple, before age 65, the man was employed full time and the woman part time</i>							
Man	16.35 (0.62)	0.19 (0.04)	0.05 (0.05)	-0.85 (0.18)	-0.22 (0.21)	1.72 (0.39)	0.50 (0.66)
Woman	20.21 (1.01)	0.18 (0.05)	0.16 (0.04)	-0.79 (0.23)	-0.71 (0.17)	1.62 (0.47)	1.53 (0.35)

* Tables 5 and A2 (appendix A) provide details on the income classifications. Standard errors are in parentheses.