Individual Income and Remaining Life Expectancy at the Statutory Retirement Age of 65 in the Netherlands

September 9, 2009

Adriaan Kalwij[a d], Rob Alessie[b d] and Marike Knoef[c d],

a) Utrecht University

b) University of Groningen and Tinbergen Institute

c) CentERdata and Tilburg University

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

Correspondence to

Adriaan Kalwij

Utrecht University

Utrecht School of Economics

Janskerkhof 12

3512 BL Utrecht

The Netherlands

Tel.: +31-30-2539664

Fax: +31-30-2537373

Email: a.s.kalwij@uu.nl

Abstract

This paper examines the association between income and remaining life expectancy at the statutory retirement age of 65 the Netherlands. A mortality risk model which explicitly controls for unobserved individual specific heterogeneity is estimated using a large administrative dataset over the period 1996-2007. Remaining life expectancy is estimated to be 16.0 years for men and 18.4 years for women in 2005. For men and women, life expectancy is about three years less for lower income than for higher income individuals. The remaining life expectancy for women with a lower income spouse is about two years less than for women with a higher income spouse. For men this difference is insignificant. An economic implication of these results is that individuals' internal rate of return of a uniform priced pension plan is positively associated with income since higher income individuals live, on average, longer.

JEL codes: C33, I10, J00

Keywords: Panel data, 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). These socioeconomic inequalities in mortality are commonly referred to as differential mortality and strongly support an inverse relationship between income and mortality risk. Interest by policymakers in this relationship stems from, among other things, that differential mortality implies a relatively bad pension deal for lower income individuals. Compared to higher income individuals, lower income individuals' internal rate of return of a uniform priced pension plan is, on average, lower since they have lower life expectancy, hence are expected to claim pension benefits for a relatively shorter period (Menchik, 1993, Brown, 2000, Simonovits, 2006, Bonenkamp, 2007, Hári, 2007). This is in particular a concern for policymakers in developed countries such as the Netherlands because a very large part of retirement consumption is financed out of pension income (Knoef et al., 2009). Furthermore, public pension policy in the Netherlands aims at redistributing income from the financially better to the financially worse off individuals and this redistribution may be adversely affected by differential mortality (Nelissen, 1999).

To this end, we empirically quantify the association between individuals' income and remaining life expectancy at the statutory retirement age of 65 in the

Netherlands.[1] The theoretical framework for much of the empirical work in economics on the relationships between socioeconomic position, health outcomes and longevity is a seminal paper by Grossman (1972). Grossman's model assumes that at the start of the planning period individuals are endowed with a stock of knowledge capital and a stock of health which deteriorates with age. Consumption is assumed to be affected by knowledge capital, e.g. through investments in education, and by the stock of health through its impact on the time spend on market activities. Individuals have an incentive to invest in the stock of health as they derive utility from the stock of health and consumption. Efficiency of the investments in the stock of health is assumed to increase with knowledge capital. These mechanisms yield that the stock of health deteriorates with age at a relatively slower rate for individuals with a higher socioeconomic position (e.g. level of education). When the stock of health falls below a threshold value the individual will decease.[2] In Grossman's model the socioeconomic position of an individual is positively related to both lifetime income and longevity which implies an inverse association between lifetime income and mortality risk. In line with this implication, and as mentioned above, empirical studies have identified an inverse relationship between individuals' financial situation and mortality risk. Estimates of this relationship for different target populations with respect to countries and age ranges and different approximations of lifetime income, show that the ratio of mortality risk of individuals in the lowest quartile of the income

[1] Previous papers have been concerned with issues such as the implications of differential mortality for the age-wealth profile (Atkinson, 1971, Shorrocks, 1975) and the incentives for saving during 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, and Van den Berg et al., 2006); however, such analyses are not an aim of this paper.

[2] Grossman (2000) provides a discussion on this and overviews the theoretical and empirical literature.

or wealth distribution over that of individuals in the highest quartile of the distribution ranges from around two in Europe up to three for the U.S.[3] In the literature from various disciplines there is an ongoing debate on the causal interpretation of this relationship and the possible pathways through which socioeconomic position affects health status and mortality (e.g., Macintyre, 1997, Marmot et al., 1991, and Smith, 1999, Snyder and Evans, 2006). This paper takes the position that there is, or might be, an association between income and mortality risk, without assuming a causal relationship, and aims at estimating this association.

The contribution of this paper to the empirical literature on the relationship between income and mortality risk is threefold. First, we quantify the association between income and mortality risk for the Netherlands for individuals aged 65 years of age or over (the retired population). As far we know this has not been done before as previous studies mentioned above for the Netherlands have most often examined the association between education and (prime age male) mortality risk (e.g. Kunst and Mackenbach, 1994). We choose the age of 65 because this is the statutory retirement age in the Netherlands and from this age onwards all individuals receive an old age public pension independent of their earnings history[4] and, depending on their earnings history, an occupational pension (e.g. Nelissen, 1999). As mentioned above, income during retirement mainly consists of pension income which is closely related to individual's earnings history and therefore considered a good proxy for lifetime income, which is the suggested measure by the economic theory discussed above.

[3] Duleep (1986) and Attanasio and Hoynes (2000) for the U.S., 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,

[4] Eligibility and amount only depends on the years registered living in the Netherlands between the ages of 15 and 65 (2% of the full public pension benefit for each year).

Second, we make a distinction between income of the individual and income of the spouse. Previous studies have used, depending on data availability, either individual income or the sum of individual and spouse's income, i.e. household income. Grossman's model, discussed above, is on individual choice and a natural starting point of analysis is the association between individual income and mortality. In his model a spouse may improve efficiency of an individuals' investment in the stock of health and the more so the higher a partner's socioeconomic position. If so, this has the implication as an individual's income is not only associated with own mortality risk but also with that of the spouse. Third, our empirical model controls for unobserved individual specific heterogeneity. Inherent to the analysis of mortality risk is that with age the sample becomes more selective with respect to both observed and unobserved characteristics. This is referred to as dynamic selection and not controlling for this may bias the results (Cameron and Heckman, 1998, Van den Berg, 2001). Panel data or duration data is required to control for dynamic selection. Most of the previous papers discussed above use (pooled) cross-section data. Using duration data Hupfeld (2009) does not control for unobserved heterogeneity and Van den Berg et al. (2006) mentions that controlling for unobserved heterogeneity does not affect the estimates of the impact of economic conditions early in life on individual mortality.

This paper is organized as follows. Section 2 describes the data. Section 3 formulates the empirical model for analysing mortality risk and discusses the estimation procedure. Section 4 reports and discusses the estimation results. Section 5 summarizes the main results and concludes.

2. The Data

Data are taken from the 1996-2007 Income Panel Study of the Netherlands (IPO, Inkomens Panel Onderzoek, CBS 2009a) and 1997-2008 Causes of Death registry (DO, Doodsoorzaken, CBS 2009b). These data are gathered by Statistics Netherlands.

IPO is a representative sample of the Dutch population. The data are an administrative panel dataset of about 92,000 individuals in 1996 and increasing to about 99,000 individuals in 2007 due to population growth. Sampling is based on the national security number of individuals. A selected individual is followed over time as long as he or she is a resident in the Netherlands on the 31st of December of the sample year. Individuals who are living in old age institutions such as nursing homes are included.[5] An individual enters the panel for the first time in the year of birth when born in the Netherlands or in the year of arrival when immigrated to the Netherlands. An individual exits the panel when emigrating from the Netherlands or when he or she has deceased. There is no panel attrition for reasons other than mortality and emigration.[6]

IPO has information on the demographic characteristics and income of each member of the individual's household. This information is obtained from official institutions and mainly the population registry and tax office. DO data provide information on the date and causes of death of all residents that have deceased during the period 1997-2008. These data are medical records provided by medical examiners

[5] Survey data used in most previous studies usually do not include the institutionalized population.

[6] Attrition in survey data is often at an annual rate of 20% or more (e.g., Attanasio and Emmerson, 2003), which may lead to inconsistent estimates of the association between income and mortality risk if attrition is, for instance, strongly related to health status.

who are legally obliged to submit these to Statistics Netherlands. The DO data also contain a person identifier that makes it possible to identify whether or not an individual in the IPO has deceased next calendar year.

We select individuals aged 65 or over and these constitute about 12.8% of the sample in 1996 and increasing to 13.9% in 2007. This raw dataset consists of 21,159 individuals who make up for 151,120 observations over the period 1996-2007. 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). This removal affects relatively more men than women (11% versus 4%). After controlling for gender, the mortality rate among the individuals who are excluded is not significantly different from the mortality rate among those included.[7] Panel attrition for reasons other than mortality, i.e. due to emigration or missing values, is about 0.3% per year. The resulting sample consists of 19,258 individuals of whom 11,601 are women and 7,657 are men. Of these individuals, 8157 (42%) are followed from the age of 65 onwards. In total we have 141,725 observations.

2.1 Variable Definitions and Descriptive Statistics

In the analysis we use the variables gender, age marital status and income. Age is defined as the age at 1st of January of each year. This date is chosen because the income is measured over a calendar year - which is the fiscal year in the Netherlands - and in this way we ensure that income at age 65 is measured over the first entire calendar year in retirement. Table 1 reports the number of observations by age and gender. The variable marital status distinguishes between a single adult household which includes divorcees (hereafter “single”), a married or cohabiting couple

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

("married") and a widowed individual. Table 1 reports marital status by age and gender. The statistics in table 1 are dominated by the well-known fact that women live, on average, longer than men. This yields, for instance, an increasing proportion of women in the sample with age (last column panel A) and, at a given age, relatively more widowed and fewer married women than men (panel B). Statistics Netherlands requires that in the tables that follow we report '-' if the number of observations on which the statistic is based is below 25.

The income data in the IPO are mainly based on tax records and contain detailed and accurate information on all income components. Income is gross of income tax and social insurance contributions and measured in 2005 euro using the consumer price index. Individual income is defined as the sum of pension, labour, transfer and capital income. Table A1, appendix A, shows more details on these components, its definitions and that, as stated in the introduction, for both men and women over 90% of income is pension income. All income components are observed at the individual level and, in case of a couple, are also observed for the spouse. Income from other household members is excluded from the analysis.[8]

Table 2 reports the distribution of income by gender and marital status. 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 distribution of income for married men, and also for singles, shows a decrease in median income with age, while mean income remains relatively constant. This is most likely the result of changing distribution of income over the cohorts (Knoef et al., 2009). The income distribution for married women is rather compressed. This is in part because many retired married

[8] About 8% of households have other household members and these are mostly children.

women have no earnings history[9] and receive only old age public pension benefits but is also because before 1990 part-time work often did not come with a pension plan or with a pension plan with a relatively high threshold before any pension contributions could be made.

A comparison of the tables for married men and women shows that women's income account for, on average, one-third of household income. Not shown here is that the rank correlation between income of the individual and spouse is about 0.10.

2.2 Differential Mortality

Mortality is defined as having deceased next year. About 38% of the individuals have deceased over the sample period. Table 3 shows the well-known patterns that mortality risk increases with age and that men have a higher mortality risk than women. Furthermore, age-specific mortality risk is lower among married individuals than among singles or widowed individuals. The sample statistics on mortality risk by gender and age compare favourably to the population statistics reported in the column "HMD", i.e. differences between the two by age and gender are small. Table 4 reports on the mortality risk by quartiles of the distribution of individual and spouse's income. Age-specific mortality risk between singles and widowed individuals are not that much different and we therefore pool them in panel A. Panel A shows that mortality risk decreases with income quartile for single men and women and this increase appears strongest up to the third quartile. A measure of differential mortality is the ratio is mortality risk among individuals in the first quartile of the income distribution over mortality risk among individuals in the fourth quartile of the income

[9] In our sample, about 25% of women aged 65-69 have an occupational pension, for women aged 70-74 this is about 20% and for women 75 or over this is about 16%.

distribution (columns "Q1/Q4"). A comparison of panels A and B shows that differential mortality is stronger for married men than for single men (2.4 versus 1.7) and for married women weaker than for single women (1.4 versus 1.9). Overall, differential mortality appears to decline with age. These statistics are in line with findings in other European studies and lower that those reported for the U.S. (see introduction). Panel C shows differential mortality with respect to spouse's income. Interestingly, for both men and women a similar differential mortality pattern is observed as when considering individual income, albeit much weaker for men. We are not aware of any comparable statistics in the literature.

3. Mortality Risk Model

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

$$H_{a+1} = \alpha_a + X_a\beta + \Lambda + \varepsilon_a , \tag{1}$$

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

In the context of this paper, and in line with Grossman's model, the latent variable H_{a+1} can be thought of as individual's stock of health. If the stock of health falls next year below a threshold, normalised to zero in the equation (1), the individual will decease. For now we suppress the index for the individual. The variable M_{a+1} denotes observed mortality at age (a+1). M_{a+1} is equal to one if an individual became a years old and has deceased at age (a+1), and zero otherwise. α_a is an age specific intercept, X_a is a (1xk) vector of individual's observed characteristics at age a, including marital status and income, with a corresponding (kx1) parameter vector β. Λ denotes an individual's unobserved characteristic and is assumed constant over time and independent of the covariates at age 65. This assumption does not exclude dependency between the covariates and Λ at later ages. The random effect is assumed to be normally distributed with mean zero. The error term ε_a is assumed to follow a logistic distribution and is independently distributed across individuals and time with mean zero and a variance that is normalized to $\pi^2/3$.

3.1 Dynamic Selection and Initial Conditions

Inherent to a study of mortality risk over the lifecycle is that the population at risk changes with age. The model outlined above explicitly accounts for random effects and thereby allows for sample selection with age on the basis of both observed and unobserved characteristics. This is referred to as dynamic selection (Cameron and Heckman, 1998). If mortality risk is inversely related to income and the unobserved characteristic (Λ), the low income (high income) individuals with a high (low) value of Λ are more likely to survive one more year than individuals who have a low income and a low value of Λ. In this example dynamic selection results in a population at risk in which the correlation between Λ and retirement income becomes increasingly negative with age up to some age and will decrease thereafter.

When all individuals are observed from the age of 65 the model outlined above takes dynamic selection into account. However, as reported in section 2, 42% of the individuals enter the sample at age 65 and the remaining 58% enter the sample at a later age. We denote the age of an individual when first observed in the sample as age τ ($\tau \geq 65$). Dynamic selection implies that the individuals entering the sample at a later age are a selective sample with respect to retirement income and their value of Λ since they survived up to age τ. Not taking this into account would yield inconsistent estimates of α_a and β (Cameron and Heckman, 1998, Van den Berg, 2001). A solution is to explicitly take the dependency between random effect and retirement income at the age of entry (after age 65) into account. Essentially the dependency between the covariates and the random effect Λ at the age of first observation τ is parametrised for individuals who enter the sample after the statutory retirement age of 65.[10] This parameterization is formalized as follows

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

where $\tilde{X}_\tau = (1, X_\tau) \times (\tau - 65)$ with a corresponding ((*k*+1)x1) parameter vector γ. Scaling the effects of the covariates at age τ with the factor $(\tau - 65)$ takes into account that the dependency between the random effect and retirement income becomes more negative with age - as discussed in the example above. θ is a random effect that is assumed independent of X_τ and normally distributed with mean zero and variance σ^2. Consistent estimates of the α_a's and β are obtained under the additional assumption formalized in equation (2).

As mentioned above, at some age the dependency between the covariates and the random effect Λ may start to decrease in absolute terms as the sample becomes more homogeneous. This would imply an inclusion in Eq.(2) of the term $\tilde{X}_\tau^{\,2} = (1, X_\tau) \times (\tau - 65)^2$. We have experimented with this and found that such a more flexible specification of Eq.(2) does not alter the main conclusions of this paper.

---

[10] This can be considered a solution in the spirit of Wooldridge (1995) who explicitly models the dependency between random effects and explanatory variables in nonlinear models.

3.2 Estimation, identification and empirical specification

Given the model outlined above, mortality risk evolves over the lifecycle conditional on observed and unobserved characteristics as follows

$$\Pr\left(M_{a+1}=1\mid X_{\tau},X_{a},M_{a}=0\right)=F\left(\alpha_{a}+X_{a}\beta+\tilde{X}_{\tau}\gamma+\theta\right). \tag{3}$$

The condition $M_a = 0$ formalizes that the population at risk are all individuals who are alive at age a. $F(.)$ is the logistic cumulative distribution function. Equation (3) is used to set up the likelihood function in which we integrate out the unobserved individual specific effect. The individual is denoted by i, $\tau(i)$ 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 has deceased at age $A(i)+1$, and zero otherwise. Maximum likelihood estimates of the model parameters are given by

$$\left(\hat{\alpha},\hat{\beta},\hat{\gamma},\hat{\sigma}\right)=$$

$$\underset{\alpha,\beta,\gamma,\sigma}{\operatorname{argmax}}\sum_{i=1}^{N}\log\left(\int_{-\infty}^{+\infty}\left(\prod_{a=\tau(i)}^{A(i)-1}\left(1-F\left(\alpha_{a}+X_{a}(i)\beta+\tilde{X}_{\tau(i)}(i)\gamma+\theta(i)\right)\right)\right)^{I(A(i)>\tau(i))}\right.$$

$$\times\left(1-F\left(\alpha_{a}+X_{A(i)}(i)\beta+\tilde{X}_{\tau(i)}(i)\gamma+\theta(i)\right)\right)^{1-m(i)}$$

$$\left.\times\left(F\left(\alpha_{a}+X_{A(i)}(i)\beta+\tilde{X}_{\tau(i)}(i)\gamma+\theta(i)\right)\right)^{m(i)}d\Phi\left(\frac{\theta(i)}{\sigma}\right)\right), \tag{4}$$

Where $\alpha=(\alpha_{65},..,\alpha_{T})$, T is the maximum age an individual may reach, 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).[11] From the point of view of a discrete hazard rate model, equation (4) imposes proportionality between the age pattern, the covariates and the random effects to ensure identification of the distribution of the random effects (see, e.g., Cameron and Trivedi 2005).

In order to estimate the association between individual's and spouse's income and mortality risk we parametrise equation (1) as follows:

$$H_{a+1} = \alpha_0 + \alpha_1 \text{AGE}_a + \beta_1 \text{MARRIED}_a + \beta_2 \text{WIDOW}_a$$
$$+\beta_3 \ln\left(\text{Y}_a^I\right) + \beta_4 \ln\left(\text{Y}_a^P\right) \times \text{MARRIED}_a + \Lambda + \varepsilon_a \qquad (5)$$

The association between mortality risk and individual retirement income (Y_a^I) and mortality risk and spouse's retirement income (Y_a^P) are given by, respectively, β_3 and β_4. We follow Snyder and Evans (2006) and use a logarithmic specification to capture the nonlinear relationship between income and mortality risk. MARRIED is a dummy variable equal to one if the individual is married or cohabiting 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. We have tried more flexible empirical specifications, for instance age specific intercepts instead of a linear function in age, and briefly discuss these in the next section. The model is estimated separately for men and women.

In addition we control in the empirical analysis for time effects. This implies that we do not explicitly model cohort effects. This might affect the estimated age pattern since an individual's birth cohort is equal to the year of observation minus

[11] The model is estimated using the *xtlogit* command of the software package STATA (www.stata.com).

individual's age. However, if, for instance, the cohort effect is linear, this will not affect the estimate of the other parameters of eq.(5).

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 this is associated with individual and spouse's income by gender and marital status at the age of 65. We refer to appendix C for the technical details.

Income during retirement depends on lifetime earnings. The way they relate depends on the rules of the public and private pension systems. We take these rules into account when calculating income during retirement conditional on income before retirement, which is referred to as pension related gross yearly salary. Income during retirement, in turn, is input of the simulation exercise. Table 5 presents for several types of households the income during retirement conditional on pension related gross yearly salary. The pension related gross yearly salary determines the occupational pension income an individual receives on top of an old age public pension. All individuals are assumed to have lived in the Netherlands from the age of 15 onwards and are therefore entitled to a flat rate public pension benefit that only depends on household composition. The spouse's age, in case of a marriage, is the same as the age of the individual. The baseline situations are for individuals with a pension related gross yearly salary equal to median income in the Netherlands in 2005, which is €29,500. We assume that occupational pension is based on 40 years of employment.

For the cohorts included in the analysis most men have worked fulltime before the age of 65 and most women have not worked and about a quarter have worked part-time before the age of 65. The presented baseline situations in table 5 take these scenarios into account.

The rules we apply for calculating income during retirement are in line with the rules as applied by the largest pension funds (see footnote table 5). For instance, these calculations 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 we consider differences from the baseline situations due to changes in the pension related gross yearly salary. 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 working life and high income refers to a pension related gross salary based on two times median income (see table A2, appendix A). These income classifications roughly correspond to the averages in the lowest and highest income quartiles for the different household types.

4. Empirical Results

The estimation results of the model outlined in section 3 are reported in table 6 and discussed in section 4.1. Tables 7 and 8 report and section 4.2 discusses the simulation results. In the following discussion, we use a 5% level of significance unless stated otherwise.

Before proceeding to the estimation results we briefly discuss model selection. We have considered models containing a nonparametric specification of the age dependency of mortality risk and quadratic log-income effects. The results of these models are in tables A3 and A4, appendix A. Statistical test results reveal for both men and women that quadratic log-income terms are jointly insignificant and, moreover, that the nonparametric age function can be restricted to a function linear in age.[12] The restricted specification is as outlined in section 3 (model 3 in table A3 and model 7 in table A4) and the results concerning the main parameters of interest are also reported in Table 6.

4.1 Estimation Results

The results in table 6 show that individual income is negatively associated with mortality risk for both men and women and the parameter estimate is roughly of the same magnitude. A statistical test does not reject the null hypothesis of equality of the association between individual income and mortality risk for men and women.[13]

Mortality risk is negatively associated with spouse's income. The association is also roughly of same magnitude for men and women but for men this association is insignificant. For women we reject a one-sided hypothesis that spouse's income has no association, against the alternative of a negative association, with mortality risk.[14] The estimates of the effects of marital status cannot be directly interpreted due to the interaction with spouse's income and we return to this in the next section.

The estimate of the standard deviation of the random effect is significant for both men and women and underlines the importance of controlling for dynamic

[12] Furthermore we have considered an interaction terms between individual income and WIDOW and MARRIED. These interaction terms turned out to be jointly insignificant for both men and women.

[13] The p-value corresponding to this null-hypothesis is 0.816.

[14] The p-value corresponding to H_0: $\beta_4<0$ is 0.040.

selection. Furthermore, as reported at the bottom of table 6, the additional covariates to control for possible dependency between the random effect and the covariates at age τ, as formulated in Eq.(2), are jointly significant for men but for women only at a 10% level of significance. We discuss the quantitative importance of random effects for differential mortality in the next section.

4.2 Simulation results

The results given in table 7, which are based on Monte Carlo simulation as outlined in section 3, quantify the association between income and remaining life expectancy at the age of 65 by gender for the different household types and income scenarios (see tables 5 and A2). The point estimates of (unconditional) remaining life expectancy at age 65 are 16.0 years for men and 18.4 for women.[15]

As discussed in section 3, the baseline situation is based on median income and considers three types of households. For single men remaining life expectancy is about 12 years and for single women about 17.5 years (first column, first to rows, table 7). Remaining life expectancy at 65 is considerably higher for married individuals, about 16 years for married men and 20.5 years for married women (first column, last to rows). The differences between single and married individuals are significant and about 4 years for men and 3 years for women.[16]

The second and third columns report for each of the three household types on the change in life expectancy associated with a 10% higher than median income for either men or women. Table 7 shows that this association is about the same across

[15] For this purpose we estimate the model with only age and year dummies. The estimates, based on simulations, for remaining life expectancy are 15.99 (0.49) for men and 18.41 (0.38) for women.

[16] Estimates of these differences are 3.957 (0.487) for men and 3.099 (0.641) for women.

gender and household type: the point estimates vary between 0.2 and 0.26 with standard errors of 0.04. Compared to individuals with median income, life expectancy at age 65 of individuals with a 10% above median income is about 2.5 to 3 months higher.

The fourth and fifth columns report on the differences in remaining life expectancies between lower income and median income individuals. An individual on minimum wage or with no earnings (low income) during working life receives only a public old age pension benefits during retirement. For single men and women the difference in remaining life expectancy is about one year (-1.20 for men, -1.06 for women). For married men this difference is about 1.6 years (third and fifth row) and for married woman about 1 year (sixth row). The sixth and seventh columns report on the differences in remaining life expectancies between higher income individuals, defined as two times median income, and median income individuals. For both men and women the point estimate of this difference is around 1.6-1.8 years (1.59 for a single man, 1.65 for a single woman, 1.78 and 1.80 for married men, and 1.60 for a married woman). To summarize, the difference in remaining life expectancy at age 65 between lower income individuals who only have a public old age pension and higher income individuals is about three years for both men and women.

Next we turn to the association between remaining life expectancy at 65 and spouse's income. Table 7, fourth and sixth rows, shows that, compared to married women with a spouses with median income, life expectancy at age 65 of women with a spouses with a 10% above median income is 0.17-0.19 years higher (almost two months). For married men (fifth line) this difference is smaller and insignificant (0.06). Turning to the last rows in columns four and six, the difference in remaining life expectancy at 65 between women with a spouse who has a low income and women with a spouse with a high income is about two years (1.26 – (-1.02)). For men this difference is less than one year and insignificant.

Finally we turn to the importance of dynamic selection. As discussed above, we find a significant presence of unobserved individual specific heterogeneity (see table 6). To quantify its importance we have estimated the model without controlling for unobserved individual specific heterogeneity. These estimation results are reported in tables A3 (model 4) and A4 (model 8) in appendix A. Next we have performed the same simulations of remaining life expectancies at age 65 as above and these results are reported in table 8. Given the standard errors, the differences with the results in table 7 are mostly insignificant. One reason for this is that although the estimated standard deviations of the random effects are significant they are relatively small compared to the standard deviation of the error term in Eq.(1) which is equal to $\sqrt{\pi^2 / 3} \approx 1.81$. Nevertheless, the overall picture that emerges from a comparison of the results in tables 7 and 8 is that the association with individual income becomes less strong when not controlling for dynamic selection. This suggests that failing to control for dynamic selection may yield an attenuation bias in the association between income and mortality risk.

5. Conclusion and Discussion

This paper examines the association between individuals' retirement income and remaining life expectancy at the statutory retirement age of 65 in the Netherlands. For this purpose a mortality risk model is estimated which explicitly controls for unobserved individual specific heterogeneity. The data are administrative data taken from the 1996-2007 Income Panel Study of the Netherlands supplemented with data from the Causes of Death registry. For the whole sample, remaining life expectancy at 65 is estimated to be about 16.0 years for men and 18.4 years for women. The main empirical findings are summarized as follows.

For both men and women we find that mortality risk is negatively associated with individual income. A 10% higher than median individual income is associated with a 2.5 to 3 months higher remaining life expectancy at 65 for both men and women. The difference in remaining life expectancy at age 65 between lower income individuals who only have a public old age pension and higher income individuals, defined as two times median income, is about three years for both men and women. A strong implication of these findings that has not yet been reported in the literature is that differential mortality with respect to individual income is equally strong for women as it is for men.

Remaining life expectancy at 65 for women is found to be strongly associated with income of the spouse. Remaining life expectancy is about two years less for women with a lower income spouse than for women with a higher income spouse. There are two reasons for this difference. Firstly, income of the spouse is negatively associated women's mortality risk (see table 6). Secondly, marriage is negatively associated with mortality risk and women benefit, on average, from an extended duration of marriage when married to a higher income man. For men we find not such associations, mainly because men benefit relatively less from a higher income spouse as they live, on average, shorter than women.

A methodological finding is a significant presence of unobserved individual specific heterogeneity for both men and women. This underlines the importance of controlling for dynamic selection and implies that ignoring dynamic selection may yield an attenuation bias in the association between remaining life expectancy at age 65 and individual income.

An important economic implication of our results, in line with previous studies, is that individuals' internal rate of return of a uniform priced pension scheme is positively associated with income since higher income individuals live, on average, longer. In other words, lower income individuals receive a bad pension deal and higher income individuals receive a good pension deal. We have quantified the association between income and remaining life expectancy at age 65 and this may be useful information for policymakers in the Netherlands who are currently debating pension reforms. For instance, in line with recent pension reforms in several European countries, the Dutch government plans to increase the statutory retirement age from 65 to 67 for public old age and occupational pensions and hereby keeping the benefits unchanged (CPB, 2009). As argued by opponents of this policy reform (FNV, 2009), this increase will have a relatively larger negative impact on the expected pension benefit duration of the lower than of the higher income individuals. The empirical results of this paper are in support of this argument and imply that the decrease in the expected pension benefit duration is about 20-25% larger for lower than for higher income individuals.[17] This is, however, only one element of the complex nature of the effects this pension reform may have on the economy as a whole and on the welfare of individuals.

Acknowledgments

We wish to thank Pierre Carl Michaud and the seminar participants at the Utrecht School of Economics, Utrecht University, and at a workshop organized by the Network for Studies on Pensions, Ageing and Retirement (NETSPAR), for valuable comments and discussions. Financial support has also been provided by Stichting Instituut GAK through NETSPAR.

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, *The Journal of Human Resources*, Vol.35, No. 1, pp. 1-29

[17] Expected benefit duration at age 65 is shortened with two years for all individuals. The decrease in expected duration is about 12.5% for men (2/16) and 11% for women (2/18.4). Expected benefit duration for lower (higher) income men is 14.5 (17.5) years, hence an expected reduction of 14% (11%), and for lower (higher) income women is 16.9 (19.9) years, hence an expected reduction of 12% (10%).

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 P.K.Trivedi. 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

CPB Netherlands Bureau for Economic Policy Analysis, 2009, Houdbaarheidseffect verhoging AOW-leeftijd, *CPB notitie*, June 19

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

FNV, 2009, Samen werken voor de AOW, Stichting FNV pers.

Gaudecker, von HM., and R.D. Scholz, 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, Anderson, 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

Knoef, M., Alessie, R.J.M., and A. Kalwij, 2009, Changes in the income distribution of the Dutch elderly over the years 1989-2020: a microsimulation, *mimeo*, Tilburg University

Kunst, A.E., and J.P Mackenbach, 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 W.D. Kelton. 1982. *Simulation Modelling and Analysis*. New York: McGraw-Hill.

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

Menchik, P.L., 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., Grønbæk, M., Christensen, U., Due, P., 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

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., 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

Van den Berg, GJ (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

Wooldridge, J. M., 1995, Selection corrections for panel data models under conditional mean independence assumptions. *Journal of Econometrics 68*, 115–132

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

Appendix A

Table A1 Components of individual income during retirement [a)]

Male income components

Age	public pension %	Occupational pension %	Labour 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 %	Labour 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 onwards receive a public old age pension. Occupational pensions are mainly pensions related to employment before age 65. Private annuities are included and may be more common among the self-employed. Labour income includes work related earnings and income from self-employment. Transfer income includes mainly alimony payments or receipts and rental subsidies. Capital income includes mainly 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.

Table A2 Income scenarios during retirement for different types of households.

	Baseline situation	Difference from the baseline situation due to differences in income					
	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	**€**	**€**	**€**	**€**	**€**	**€**	**€**
Man (fulltime)	29500	32450		16392		59000	
Woman (part-time)	14750		16225		8196		29500
Annual income during retirement by household type							
A single person household							
Man or Woman, were fulltime employed before age 65	20650	22715	22715	11705	11705	41300	41300
Woman, was part-time employed before age 65	16178		17210		11705		26503
A two person household, before age 65 the man was fulltime employed 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	
A two person household, before age 65 the man was fulltime employed 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	
	Parameter estimate	*Standard error*	*Parameter estimate*	*Standard error*	*Parameter estimate*	*Standard error*	*Parameter estimate*	*Standard error*
Constant	-1.854	0.485	-2.597	0.248	-11.385	1.016	-9.852	0.272
Year = 1996	0.000		0.000		0.000		0.000	
Year = 1997	0.045	0.095	0.046	0.095	0.037	0.096	0.050	0.094
Year = 1998	0.107	0.098	0.110	0.098	0.090	0.099	0.114	0.093
Year = 1999	0.047	0.104	0.053	0.104	0.029	0.105	0.064	0.094
Year = 2000	0.036	0.111	0.045	0.111	0.009	0.111	0.055	0.094
Year = 2001	0.094	0.118	0.108	0.118	0.059	0.117	0.118	0.093
Year = 2002	0.040	0.128	0.056	0.128	-0.002	0.125	0.068	0.093
Year = 2003	-0.029	0.138	-0.011	0.137	-0.079	0.133	0.005	0.095
Year = 2004	-0.047	0.147	-0.026	0.146	-0.108	0.141	-0.011	0.094
Year = 2005	-0.198	0.157	-0.174	0.157	-0.272	0.150	-0.161	0.097
Year = 2006	-0.186	0.166	-0.161	0.166	-0.270	0.156	-0.144	0.096
Year = 2007	-0.225	0.176	-0.198	0.176	-0.317	0.165	-0.176	0.097
Age specific dummy variables [a)]	included		included		excluded		excluded	
Age					0.139	0.015	0.111	0.003
Single	0.000		0.000		0.000		0.000	
Married	-0.406	0.550	-0.556	0.259	-0.551	0.266	-0.183	0.223
Widowed	-0.324	0.092	-0.321	0.093	-0.325	0.096	-0.217	0.069
$Ln(Y^I)$	-0.893	0.295	-0.396	0.057	-0.405	0.058	-0.313	0.041
$Ln(Y^P)$	-0.233	0.428	-0.105	0.112	-0.117	0.114	-0.196	0.100
$Ln(Y^I)^2$	0.083	0.049						
$Ln(Y^P)^2$	0.024	0.086						
Age x (τ-65)	0.000	0.056	-0.050	0.024	-0.071	0.022		
Single x (τ-65)	0.000		0.000		0.000		0.000	
Married x (τ-65)	0.069	0.045	0.046	0.023	0.045	0.024		
Widowed x (τ-65)	0.016	0.010	0.015	0.010	0.014	0.010		
$Ln(Y^I)$ x (τ-65)	-0.026	0.036	0.010	0.005	0.010	0.005		
$Ln(Y^P)$ x (τ-65)	-0.031	0.035	-0.011	0.010	-0.010	0.011		
$Ln(Y^I)^2$ x (τ-65)	0.006	0.006						
$Ln(Y^P)^2$ x (τ-65)	0.004	0.007						
Stand.dev. random effect	0.246	0.128	0.250	0.156	0.420	0.155		
Log-likelihood value	-10888.2		-10892.0		-10906.5		-10915.2	
Number of observations	54617		54617		54617		54617	
Number of individuals	7657		7657		7657		7657	
Number of parameters	57		53		23		17	

	p-value
LR-test, model 2 against model 1	0.102
LR-test, model 3 against model 2	0.518
LR-test, model 4 against model 3	0.008

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

Table A4 Estimation results for women

Mortality Risk	Model 5		Model 6		Model 7		Model 8	
	Parameter estimate	*Standard error*	*Parameter estimate*	*Standard error*	*Parameter estimate*	*Standard error*	*Parameter estimate*	*Standard error*
Constant	-2.666	0.654	-3.192	0.276	-10.487	0.846	-11.381	0.256
Year = 1996	0.000		0.000		0.000		0.000	
Year = 1997	0.119	0.086	0.118	0.086	0.137	0.085	0.119	0.084
Year = 1998	0.137	0.090	0.134	0.090	0.168	0.088	0.134	0.084
Year = 1999	0.202	0.096	0.198	0.096	0.247	0.091	0.197	0.082
Year = 2000	0.165	0.106	0.161	0.106	0.226	0.098	0.160	0.083
Year = 2001	0.183	0.116	0.179	0.116	0.260	0.105	0.181	0.082
Year = 2002	0.221	0.127	0.217	0.127	0.311	0.112	0.216	0.082
Year = 2003	0.122	0.140	0.118	0.140	0.230	0.122	0.119	0.084
Year = 2004	0.089	0.152	0.086	0.152	0.213	0.131	0.088	0.084
Year = 2005	0.089	0.164	0.086	0.164	0.227	0.139	0.090	0.083
Year = 2006	-0.129	0.178	-0.131	0.178	0.025	0.150	-0.125	0.087
Year = 2007	0.096	0.188	0.093	0.188	0.264	0.157	0.100	0.083
Age specific dummy variables [a)]	included		included		excluded		excluded	
Age					0.110	0.013	0.121	0.002
Single	0.000		0.000		0.000		0.000	
Married	-0.952	0.740	-0.348	0.228	-0.349	0.228	-0.091	0.185
Widowed	-0.092	0.085	-0.090	0.085	-0.087	0.085	-0.036	0.050
$Ln(Y^I)$	-0.812	0.425	-0.456	0.072	-0.454	0.072	-0.377	0.050
$Ln(Y^P)$	0.256	0.482	-0.130	0.074	-0.129	0.074	-0.170	0.060
$Ln(Y^I)^2$	0.059	0.072						
$Ln(Y^P)^2$	-0.064	0.078						
Age x (τ-65)	0.021	0.048	-0.029	0.023	-0.012	0.020		
Single x (τ-65)	0.000		0.000		0.000		0.000	
Married x (τ-65)	0.059	0.057	0.032	0.018	0.033	0.018		
Widowed x (τ-65)	0.005	0.006	0.004	0.006	0.004	0.006		
$Ln(Y^I)$ x (τ-65)	-0.026	0.030	0.008	0.005	0.008	0.005		
$Ln(Y^P)$ x (τ-65)	-0.028	0.039	-0.006	0.006	-0.006	0.006		
$Ln(Y^I)^2$ x (τ-65)	0.006	0.005						
$Ln(Y^P)^2$ x (τ-65)	0.004	0.007						
Stand.dev. random effect	0.254	0.152	0.252	0.134	0.233	0.046		
Log-likelihood value	-14804.7		- 14807.6		-14824.3		- 14829.2	
Number of observations	87108		87108		87108		87108	
Number of individuals	11601		11601		11601		11601	
Number of parameters	57		53		23		17	
		p-value						
LR-test, model 6 against model 5		0.215						
LR-test, model 7 against model 6		0.299						
LR-test, model 8 against model 7		0.137						

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

Appendix B

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

$$\Pr\left(D = d \mid D \geq \tau; X, \theta\right)$$

$$= \frac{\Pr\left(D = d \cap D \geq \tau \mid x, \theta\right)}{\Pr\left(D \geq \tau \mid x, \theta\right)} = \frac{\Pr\left(D = d \mid x, \theta\right)}{\Pr\left(D \geq \tau \mid x, \theta\right)},$$

$$\text{,,} \qquad = \frac{\Pr\left(D \geq d \mid x, \theta\right)}{\Pr\left(D \geq \tau \mid x, \theta\right)} \times \frac{\Pr\left(D = d \mid x, \theta\right)}{\Pr\left(D \geq d \mid x, \theta\right)},$$

$$\text{,,} \qquad = \frac{\Pr\left(D \geq d \mid x, \theta\right)}{\Pr\left(D \geq \tau \mid x, \theta\right)} \times \Pr\left(D = d \mid D \geq d; X, \theta\right),$$

$$\text{,,} \qquad = \frac{\Pr\left(D \geq d \mid x, \theta\right)}{\Pr\left(D \geq \tau \mid x, \theta\right)} \times F\left(\alpha_a + X_d\beta + \tilde{X}_\tau\gamma + \theta\right),$$

$$\text{,,} \qquad = \frac{\prod_{a=0}^{d-1}\left(1 - F\left(\alpha_a + X_a\beta + \tilde{X}_\tau\gamma + \theta\right)\right)}{\prod_{a=0}^{\tau-1}\left(1 - F\left(\alpha_a + X_a\beta + \tilde{X}_\tau\gamma + \theta\right)\right)} \times F\left(\alpha_a + X_d\beta + \tilde{X}_\tau\gamma + \theta\right),$$

$$\text{,,} \qquad = \prod_{a=\tau}^{d-1}\left(1 - F\left(\alpha_a + X_a\beta + \tilde{X}_\tau\gamma + \theta\right)\right) \times F\left(\alpha_a + X_d\beta + \tilde{X}_\tau\gamma + \theta\right),$$

and for $d=\tau$ given by

$$\Pr\left(D = d \mid D \geq \tau; X, \theta\right)$$

$$= \frac{\Pr\left(D = d \mid x, \theta\right)}{\Pr\left(D \geq d \mid x, \theta\right)} = \frac{\Pr\left(D \geq d \mid x, \theta\right)}{\Pr\left(D \geq d \mid x, \theta\right)} \times \Pr\left(D = d \mid D \geq d; X, \theta\right),$$

$$\text{,,} \qquad = F\left(\alpha_a + X_d\beta + \tilde{X}_\tau\gamma + \theta\right).$$

Appendix C Monte Carlo simulations

The Monte Carlo simulations are based on the distribution function of remaining life duration at age 65 as derived in appendix B. We perform the Monte Carlo simulations as follows. We define a group of reference individuals - referred to as the baseline - with the same characteristics, for instance married men with median income. The estimates of the parameter 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 the age of 65 each individual of this reference group has deceased next year. Next we compare these probabilities with random drawings from the uniform distribution to simulate whether or not each individual has deceased next year (see e.g., Law and Kelton 1982). Finally, we simulate age-by-age the mortality status for each individual in the reference group until 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 the age of 65 up to 105. These sequences are used 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 the differences between these simulation outcomes and the baseline simulation outcomes can reveal the association of this covariate with remaining life expectancy at 65, holding all other factors constant. 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 around the parameter estimates.

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 Men				Marital Status 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 The 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 columns labelled 'HMD' provide population statistics from the Human Mortality Database (2008) over the period 1996-2006.

	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 table 2.

Table 5 Income during retirement for different types of households.

Cells: yearly amounts in 2005	**€**
Statutory minimum wage, fulltime	16392
Old age public pension for a single person household	11705
Old age public pension per person for a two person household	8018
	Baseline situation
Pension related gross yearly salary, Median	**€**
Man (fulltime)	29500
Woman (part-time)	14750
Annual income during retirement by household type [a)]	
A single person household	
Man (fulltime employed before age 65)	20650
Woman (part-time employed before age 65)	16178
A two person household, before age 65 the man was fulltime employed and the woman was not employed	
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
A two person household, before age 65 the man was fulltime employed and the woman part-time	
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 pension income is based. Income during retirement is calculated using the formula Y+0.7xFTEx(Y-10/7 x FRANCHISE), a widowed individual receives 5/7 x 0.7 x FTE x (Yp-10/7 x FRANCHISE) where Yp refers to the pension related gross salary of the deceased spouse. FRANCHISE is equal to €11705 (the old age public pension received by a person living alone). FTE is equal to 1 if was fulltime employed and 0.5 if part-time.

Table 6 Estimation results [a)]

Dependent variable: Mortality Risk	Men		Women	
Covariates, parameter, Eq.(6)	*Parameter estimate*	*Standard error*	*Parameter estimate*	*Standard error*
Constant, α_0	-11.085	1.016	-10.487	0.846
Age, α_1	0.139	0.015	0.110	0.013
Single	0.000		0.000	
Married, β_1	-0.551	0.266	-0.349	0.228
Widowed, β_2	-0.325	0.096	-0.087	0.085
Ln(Individual income), β_3	-0.405	0.058	-0.454	0.072
Ln(Spouse income) x Married, β_4	-0.117	0.114	-0.129	0.074
Standard deviation random effect, σ	0.420	0.155	0.233	0.046
Log-likelihood value	-10906.5		-14824.3	
Number of parameters	23		23	
Number of observations	54617		87108	
Number of individuals	7657		11601	
	p-value		*p-value*	
F-test for joint significance of time dummies	0.034		0.007	
F-test for joint significance of initial conditions, see Eq. (2)	0.013		0.086	

a) The estimation results of the complete models are in tables A3 and A4, appendix A.

Table 7: Simulations of remaining life expectancy at the statutory retirement age of 65 (in years) by gender and household type.[a)]

	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*
A single person household							
Man	12.26 (0.60)	0.21 (0.04)		-1.20 (0.18)		1.59 (0.25)	
Woman	17.54 (0.85)		0.21 (0.04)		-1.06 (0.20)		1.65 (0.32)
A two person household, before age 65 the man was fulltime employed and the woman was not employed							
Man	15.86 (0.60)	0.26 (0.04)		-1.65 (0.24)		1.80 (0.27)	
Woman	19.57 (0.78)	0.19 (0.05)		-1.12 (0.29)		1.31 (0.37)	
A two person household, before age 65 the man was fulltime employed and the woman was part-time employed							
Man	16.22 (0.61)	0.25 (0.04)	0.06 (0.05)	-1.63 (0.24)	-0.36 (0.26)	1.78 (0.27)	0.49 (0.38)
Woman	20.64 (0.84)	0.17 (0.05)	0.20 (0.04)	-1.02 (0.26)	-1.07 (0.20)	1.26 (0.35)	1.60 (0.30)

[a)] Tables 5 and A2 provide details on the income classifications. Standard errors are in parentheses.

Table 8: Simulation results using models 4 and 8 (Tables A3 and A4), i.e. models without controlling for unobserved heterogeneity. [a)]

	Baseline situation	Difference from the baseline situation due to differences in 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*
A single person household							
Man	13.09 (0.49)	0.19 (0.03)		-1.08 (0.15)		1.41 (0.22)	
Woman	17.35 (0.45)		0.16 (0.02)		-0.84 (0.12)		0.97 (0.14)
A two person household, before age 65 the man was fulltime employed and the woman was not employed							
Man	16.16 (0.37)	0.23 (0.03)		-1.45 (0.21)		1.63 (0.23)	
Woman	18.80 (0.39)	0.17 (0.03)		-1.03 (0.20)		1.20 (0.24)	
A two person household, before age 65 the man was fulltime employed and the woman was part-time employed							
Man	16.74 (0.41)	0.23 (0.04)	0.11 (0.05)	-1.42 (0.20)	-0.58 (0.20)	1.62 (0.22)	0.86 (0.31)
Woman	19.69 (0.38)	0.16 (0.03)	0.16 (0.02)	-0.88 (0.12)	-0.88 (0.12)	1.14 (0.23)	1.28 (0.18)

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