Comparing the observed and unobserved components of the long arm of childhood – Evidence from Finnish register data on midlife mortality from siblings and their parents

Online Appendix

Appendix – Methods

Although we interpret the coefficients of our models in terms of hazard ratios from a proportional hazards model, it helps to look at the model applying the accelerated failure time (AFT) metric (Lambert et al. 2004), to understand how sibling similarity is estimated. The AFT formulation of the model is completely equivalent to the proportional hazard model, and is used only for illustrative purposes.

The formal regression equation for the frailty model on the AFT metric that we will use is defined as:

$$\log(y_{fst}) = a + b_w * t_{fs} + c * male_{fs} + b_m * t_{fs} * male_{fs} + \boldsymbol{X_{fs}\gamma} + \log(Z_f) + \log(e_{fst}) \quad (1)$$

$$\log(Z_f|X_{fs}) \sim N(0, \sigma_Z); \text{VAR}(\log(Z_f)) = \Theta$$

$$\log(e_{fs}) \sim G\left(0, \frac{\pi}{\sqrt{6}}\right);$$

he index f stands for family, s for sibling, t for the age in years above 35 and c represents the mortality difference between men and women at the beginning of the analysis period (age 35). Z_f is the frailty shared by siblings within a family. It is assumed that the random components $\log(Z_f)\ and\ \log(e_{fst})$ are uncorrelated and $\log(Z_f)$ is assumed to follow a normal distribution with a mean of zero which is a formulation more closely related to the methodological tradition of multilevel modeling than of shared frailty models (Pankratz et al. 2005). In our sensitivity analyses we also specify Gamma and inverse Gaussian distributions for $\log(Z_f)$ - which is more common in the literature on shared frailty – and discuss the (small) differences to the normality assumption The shape parameter b, representing the increase in mortality hazard per year, is allowed to differ between men ($b_w + b_m$) and women (b_w) reflecting that mortality risk increases faster for men than for women in midlife. **X** is a vector of observed predictors and $\boldsymbol{\gamma}$ the respective vector of coefficients. $\text{Log}(y_{fst})$ is the log. survival time of each sibling nested in a family. This is the accelerated failure time metric which we use for illustrational purposes to demonstrate the estimation of the sibling similarity. The error term e_{fs}follows the standard exponential distribution with a mean and variance of 1. The logarithm of a standard exponentially distributed variable is the standard extreme value distribution or standard Gumbel (G) distribution, which has a variance of $\frac{\pi^2}{6}$ (Allison 2010, p. 78). The family specific error-term is estimated as the variance of the intercept α, which is the log-baseline survival time. The metric of the family error component is therefore on the log-survival time scale and to put it into relation to this log. survival time scale it has to be compared to the variance of the standard Gumbel distribution. Following the method suggested by Goldstein et

al. (2002) we calculate the sibling similarity as an approximation of the intra-class correlation (ICC) in a linear multilevel model from the estimated variance on the family level and the assumed variance of the individual error component:

$$\hat{\rho}_f = \frac{\hat{\Theta}}{\hat{\Theta}+\sigma_{\hat{e}}^2} = \frac{\hat{\Theta}}{\hat{\Theta}+\frac{\pi^2}{6}} \quad (2)$$

A confidence interval can be derived for this statistic based on a standard error, which is derived using the delta method (Stata's nlcom command, see also (1992)).

The equivalence between our estimate of sibling similarity and the calculation of an ICC for a continuous outcome is, firstly, that the estimate is bounded by 0 and 1, and, secondly, that higher values indicate higher similarity between siblings relative to the differences between families. However, as the variance on the sibling level is not freely estimated, but fixed, we must be cautious in the interpretation of the numerical value of the sibling similarity. For low and average levels of variation between families, the value can be interpreted as an approximation of the variance in survival time (or the hazard of mortality) on the family level, compared to the total variation, which consists of family level plus individual (sibling) level. Note that an estimate of sibling similarity can also be derived from the proportional hazard model, and does not require the AFT metric, because the proportional hazard and AFT are equivalent models, differing only in the way coefficients are reported. The estimate of the shared frailty parameter is the same across the two metrics. Therefore, sibling similarity is not affected by our choice of metric.

In addition to the MHR and sibling similarity, a third statistic we use to establish the importance of familial influence is calculated by placing the estimate of familial variance in relation to the shape parameter of the Gompertz model. The shape parameter tells us how much the hazard increases per year of age. Consequently, we can divide the estimate of the square root of the familial variance by the shape parameter and make the following statement: Being in a family with one SD higher familial influence affects the mortality hazard in the same way as X years of aging. We call this equivalent years of aging (EYA):

$$\widehat{EYA} = \frac{\sqrt{\hat{\theta}}}{\hat{b}} \quad (5)$$

As with sibling similarity and MHR, a confidence interval can be derived for this ratio based on a standard error, which is derived using the delta method. While sibling similarity (as the intra-class correlation of the variance partition component) and the MHR have been used in previous research, we propose the EYA as a new approach of estimating the total familial influence on mortality. While the MHR gives us an impression of the *relative* influence of the family on mortality, the EYA provide a metric that enables us to more easily judge the *absolute* influence of the family. The EYA can be thought of as the number of years individuals from an average risk family need to age in order to reach the same level of mortality hazard as individuals from a higher risk family. Note that we will interpret the hazard rate in the context of the EYA as equal to the hazard, because – for the very low incidences of mortality in the age range under observation – the hazard rate and the hazard are almost identical. The resulting error is small, and is justified by the ease of interpretation when using this method. This way of interpreting the hazard rate is similar to treating the odds ratio as a risk ratio, under the rare disease assumption (Greenland and Thomas 1982).

While a Gompertz distributed survival time is quite common in the literature, we do compare the cumulative hazard function based on our first model, which relies on the assumption of Gumbel-distributed individual error-terms with the cumulative hazard function from the non-parametric Nelson-Aalen (NA) estimator, which is not based on said assumption. Figure A1 shows the comparison of the two estimates of the cumulative hazard function for all-cause mortality. The functions are very similar, and the parametric function lies almost exclusively within the 95% CI of the NA estimator. In Fig. A15 we show the comparison for all groups of cause-specific mortality. For all groups of causes of death the functions are very similar, and the hazard function is slightly overestimated only at higher ages (65+) for men, and for CVD related deaths. A stronger deviation is present for alcohol-related deaths above the age of 60, for which the parametric model underestimates the mortality hazard compared to the NA estimator. In sum, the simplifying assumption regarding the distribution of the error-term seems justified, and allows us to obtain an estimate of sibling similarity that approximates the intra-class correlation.

We can also use the variance estimate from our models to calculate how large the hazard ration of a variable that is standardized to have a variance of 1 on the family level needs to be to completely explain differences between families.

The restriction of the model is that the coefficient of the shared frailty parameter is 1. This assumption allows estimating the variance. By reversing the assumptions and setting the variance to 1 we get a clear relationship between the estimated variance and the respective HR it would have, if the variance would be fixed instead of the coefficient. It is the fixed coefficient (1) divided by the factor that the shared frailty parameter needs to be multiplied by to get a variance of 1, which is the ratio of 1 to the standard deviation of the shared frailty parameter:

$$HR_{fam} = \exp\left(\frac{1}{\frac{1}{\sqrt{\Theta}}}\right) = \exp(\sqrt{\Theta})$$

This is numerically close to the MHR, differing only by a factor of $\sqrt{2}$*0.6475=.9539.

Appendix – Additional results

The baseline model contains gender, cohort, a gender-specific shape factor, and a random-intercept term (shared frailty) for each family (group of siblings). Table 1 contains the estimates of individual and family-level characteristics on all-cause mortality. The variance estimate for frailty is 0.36 on the hazard scale, which translates into an estimate of sibling similarity of 0.18. We can say that almost a fifth of the variation in survival time is estimated to be between families, while 82% is between siblings.

The median hazard ratio in the baseline model is 1.77. This means that, on average, between a pair of families randomly drawn from the population, the difference in mortality risk is 77% higher in the higher risk family than in the lower risk family.

Compared to the average increase in hazard per year, being in a family that is one standard deviation below the average in survival time is approximately the same as 8.81 equivalent years of ageing (EYA averaged across gender). The family impact on the hazard rate expressed as the EYA is illustrated in Fig. A2. It plots the predictions of the hazard rate of

mortality from the shared frailty model for an average family and a family that is one standard deviation above (high risk family) and below (low risk family) the average family, respectively. The equivalent years of ageing are visually represented as thick dash-dot lines. They show that high risk men reach a hazard of 1% at age 51.88, while men from an average family first reach this hazard at age 61.64. For women, a hazard of 0.5% is reached at age 55.90 in the high risk families, and only at age 64.71 in the average risk group. Figure A3 presents the distribution of the EYA averaged across men and women. It shows the size of the high and low risk groups (+1Sd and -1SD, respectively) as a fraction of the total population (about 17% each). Furthermore, the 1^{st} and 99^{th} percentiles are indicated. Individuals from families with extremely low mortality risk are about 24.5 years behind in terms of their mortality risk, while the highest risk families (99^{th} percentile) are almost 24.5 years ahead.

In the demography model, we add variables for the age of parents at the individual's birth, differences between regions in Finland, the number of siblings in the family, and an indicator for individuals with Swedish as their mother tongue. The mortality risk of children whose mother tongue is Swedish is reduced by a factor of 0.61. There are also notable mortality differences between the regions in Finland, with Eastern Finland and the province of Uusimaa having slightly increased mortality compared to Western Finland (HR respectively: 1.13 and 1.15). There is no strong association between parents' age at birth (maternal or paternal) and mortality. The differences in mortality between number of siblings in the family are also small and non-significant. Overall, sibling similarity is not influenced notably (0.17), and nor is the MHR (1.75) or the equivalent years of aging (8.66), meaning that familial influence on mortality risk cannot be traced back to similarity of siblings with regard to language, regional parity, or parental age at birth.

The parental SEP model includes the education and occupation of the parents. A lower parental education level ("less than primary school or no information" compared to "past primary school") is associated with higher mortality (HR 1.16). We can further see that parental occupational status is also associated with midlife mortality. Compared to professionals (higher white collar), the HR for blue-collar and farm workers is 1.16; other differences are smaller and not statistically significant. The number of people living in the household per heated room is an indicator of available resources in childhood, but is only slightly and non-significantly associated with mortality. Regarding sibling similarity and

equivalent years of aging on mortality, we can see that the inclusion of parental SEP variables has little effect. Sibling similarity is still 0.17, MHR is minimally reduced to 1.73, and the equivalent years of aging are 8.48. Substantively, these changes are negligible, and we thus conclude that parental SEP has some association with mortality, but does not contribute substantially to the explanation of total familial influence on midlife mortality.

The individual SEP model adds variables for the individuals' own SEP at age 35. This includes education, income, home ownership, occupational position, and employment status. All of the dimensions of individuals' SEP exert an influence on mortality separately. For example, individuals in the lowest income decile have a mortality risk 1.99 times higher than those in the highest decile. Compared to those with higher tertiary education, individuals with only basic or unknown education have a mortality risk which is 1.57 times higher. Individuals who rent have significantly increased mortality risk compared to those who own or part-own a house at the age of 35 (HR 1.31). Lastly, compared to upper white collar workers, blue collar workers' mortality risk is 1.23 times higher.

The socioeconomic stratification variables of individuals at age 35 explain a larger proportion of sibling similarity, in addition to the small explanatory contribution of the demography and the parental SEP models. The last model reduces the conditional sibling similarity to 0.14. The median hazard ratio is 1.64. The equivalent years of aging are reduced to 7.54. We can now see that this average difference between families is larger than any difference between occupational groups (max HR 1.23) and between home owners and non-owners. It is about the same size as the difference between the lowest educational group and the highest. It is somewhat smaller than the difference between the highest and lowest income deciles, and also somewhat smaller than the difference between unemployed and employed individuals. For all-cause mortality we can conclude that, first, the average difference in mortality risk between families is almost as large as the strongest differences we find between social groups and, second, only individuals' own SEP contributes a relevant portion to the explanation of familial influences on all-cause mortality. In total, only about 20 percent of the familial influence could be explained jointly by demographic, parental, and individuals' own socioeconomic factors. As we proposed in $\boldsymbol{H_1}$, the *unobserved arm* is of greater magnitude than the *observed long arm of childhood*.

Cause-specific familial influence

In this section, we examine differences in the magnitude of sibling similarity, and the proportion of similarity explained by the demography, parental, and individual SEP models between causes of death. Table 3 in the appendix lists the relative frequency of causes of death in the sample. Figure A4 shows sibling similarity by cause of death, and Fig. A5 shows the EYA by cause of death (in addition to the MHR results in Fig. 1 in the main text).

The highest similarity is found for alcohol-related deaths (0.36), but similarity in CVD (0.33) and accidental and violent (0.25) deaths are also markedly higher than for all-cause mortality. Lung cancer (0.29) also shows higher sibling similarity than all-cause mortality. Other types of cancer show a similar total familial influence to all-cause mortality (0.19). The MHR follows the same pattern as sibling similarity. While the average difference in mortality risk between two families is about 77% for all-cause and cancer mortality, it is about 100% in accidents and violence, 120% in lung cancer, 137% in CVD, and about 150% in alcohol-related mortality. It is interesting that the EYA comparison does not follow the same picture. Lung cancer estimates are about the same size as for all-cause mortality (about 9 years). This means that if we take into account the increase in mortality risk with increasing age, the relative influence of family is the same for all-cause as for lung cancer-specific mortality. Therefore, the way in which we measure familial influence for the comparison of familial influence between causes of death does make a difference. Further, a reliable estimate for accidents and violence could not be calculated, because the shape parameter (log. increase in hazard with age) approached zero, and the values became unreasonably high, which is a common drawback of using a ratio.

Similar to the result for all-cause mortality above, parental and individuals' own SEP can only explain a small portion of the familial influence on mortality. The largest portion is explained by the influence of individuals' SEP on mortality due to lung cancer. As smoking shows a strong social gradient, individual SEP can explain an estimated 27% of the differences in lung cancer mortality between families. The cumulative explanatory power for other causes of death lies between 10% (alcohol related) to 15.41% (accidents and violence), which is smaller than the explicable familial influence on all-cause mortality. Despite the fact that we can find clear and strong social gradients in all cause-of–death groups, we can only attribute mortality differences between families to a maximum of one quarter of our measures of

social stratification. All three measures show that the differences in the level of familial influence between causes of death are much higher than the share of familial influence that can be explained by SEP (the differences between models within each cause of death), indicating that there is much more variation in the strength of the long arm of childhood across causes of death than there is between the observed and unobserved components of the arm.

Based on the MHR and sibling similarity – both measures of relative influence – we find considerable evidence to suggest that familial influence is strongest in those groups of causes that are more strongly related to behavioral factors. The pattern is less clear when we use equivalent years of aging as a measure of absolute influence. This discrepancy could be explained by the fact that relative measures do not take into account the increase in mortality hazard with age. Familial influence on CVD and lung cancer was found to be larger, but so was the increase of mortality hazard with age. In future studies, more in-depth analyses regarding familial influence on more distinct groups of causes of death would be interesting. The familial influence on lung cancer should be given special consideration, because the overall familial influence on this cause of death is larger than for all-cause mortality, and the relative importance of the visible arm (meaning the part that can be attributed to the social stratification of parents and their children) is considerably larger compared to other groups, although it remains smaller than its unobserved counterpart. This indicates that determinants of lung cancer mortality, primarily smoking (Fenelon and Preston 2012), are especially subject to observable social influences, a result that has also been found in other studies (e.g. Geyer 2008; Kulik et al. 2013; Mackenbach et al. 2004).

Appendix Sensitivity Analysis

In our analysis, the estimation of the variance term for the shared frailty parameters is the key element. In the results reported above we assumed a Gaussian distribution of the frailty parameters; however, the results might be sensitive to this assumption. Therefore, we re-ran the analyses for all-cause mortality using gamma and inverse Gaussian distribution, to

see whether the results would change substantially, due to the specification of the distribution of the frailty parameter. The gamma model for the frailty distribution has been used by numerous researchers (e.g. Manton and Stallard 1981; Vaupel and Yashin 1985). The other common distribution, inverse Gaussian, was introduced as a frailty model by Hougaard (1984). Figure A7 and A8 in the appendix show the measures of familial influence for all-cause mortality according to the distributional assumption for shared frailty parameters. We can see that a Gaussian distribution results in slightly lower overall estimates of familial influence, but the change in familial influence for the different models is proportional across specifications. We therefore take a slightly conservative estimate regarding the size of influence of shared family characteristics; nevertheless, our conclusions regarding the explanation of familial influence are not affected by the choice of distribution for the shared frailty parameters.

A second aspect that might influence the conclusions from our analyses is gender specificity in familial influence on midlife mortality. In our analysis, we combined brothers and sisters – however, both gender-specific parenting as well as gender differences in mortality (Hamil-Luker and O'Rand 2007) could lead to different results when estimating familial influence separately for brothers and sisters. To test the gender specificity of our results, we repeated our analyses for men and women separately, analyzing only those families with at least one brother and sister pairing, and dropping all families where the children are of only one gender. We also repeated the analyses for cause-specific mortality if the brother and sister subsample contained more than 500 deaths, to ensure sufficient inter-family and intra-family variation. The results of our analysis of brother and sister similarity are reported in Fig. A9-14. We can see that the differences in estimates of familial influence to the pooled sample are minor, and the explanatory contribution of the four models are very similar in size. We therefore conclude that familial influence on mortality hazard is not gender specific for our sample.

Appendix - Tables

Table A1 – Summary statistics

	Mean	SD
Demography		
Birth cohort	*1943.73*	*4.19*
Age (years) at last observation	*61.62*	*5.69*
Female	*0.49*	
Native language:		
Finnish	0.93	
Swedish	0.07	
Mother's age at birth:		
14-24	0.21	
25-35	0.58	
35+	0.19	
no valid info	0.01	
Father's age at birth:		
14-24	0.08	
25-35	0.52	
35+	0.33	
no valid info	0.06	
Region		
Western Finland	0.44	
Eastern Finland	0.38	
Lapland	0.06	
Uusimaa	0.12	
Number of siblings		
2	0.34	
3	0.27	
4	0.18	
5+	0.21	
Parental SEP		
Parental education:		
Did not go to school, unknown	0.17	
Primary school	0.73	
Past primary	0.10	
Father's occupational status		
Professional/administrative	0.14	
Workers & agriculture workers	0.41	
Farmers	0.27	
Farmers (10+ ha)	0.08	
Employer/self-employed	0.08	
Other, unknown	0.01	
Number of persons per heated room		
up to 1 person	0.05	
1-2 persons	0.37	
2-3 persons	0.28	
3 and more persons	0.29	
Unknown	0.01	
Own SEP (at age 35)		
Income		
1st decile	0.10	
2nd decile	0.08	
3rd decile	0.08	
4th decile	0.08	
5th decile	0.08	
6th decile	0.08	
7th decile	0.08	
8th decile	0.08	

9th decile	0.08
10th decile	0.09
Unknown	0.16
Education	
Basic or unknown	0.53
Upper secondary (lower track)	0.24
Upper secondary (higher track)	0.11
Lower-degree tertiary	0.07
Highest-degree level tertiary	0.05
Home ownership	
No owner	0.30
Owns house/share	0.56
Unknown	0.14
Occupation	
Self-employed	0.11
Upper white-collar	0.12
Lower white-collar	0.26
Blue-collar	0.31
Other/unknown	0.20
Employment status	
Employed	0.71
Unemployed	0.02
Homemakers	0.06
Others/Unknown	0.21

Table A2 – Relative frequency of causes of death in the sample of siblings

Cause of death	Freq.	Percent
Cancer (other than lung cancer)	2462	22.49
Cardiovascular	3188	29.12
Alcohol	1085	9.91
Accidents and Violence	1952	17.83
Lung Cancer	613	5.60
Other	1321	12.07
Unknown	327	2.99
Total	10948	100

Table A3 – Summary statistics comparing sibling sample with only children

	Sibling Sample	Only-Children
Demography		
Birth cohort	1943.73	1943.86
Age (years) at last observation	61.63	61.40
Female	0.49	0.49
Native language:		
Finnish	0.93	0.90
Swedish	0.07	0.10
Mother's age at birth:		
14-24	0.21	0.29
25-35	0.58	0.44
35+	0.19	0.25
no valid info	0.01	0.03
Father's age at birth:		
14-24	0.08	0.14
25-35	0.52	0.39
35+	0.33	0.29
no valid info	0.06	0.18
Region		

Western Finland	0.44	0.47
Eastern Finland	0.38	0.31
Lapland	0.06	0.03
Uusimaa	0.12	0.19
Number of siblings		
2	0.34	0.00
3	0.27	0.00
4	0.18	0.00
5+	0.21	0.00
Parental SEP		
Parental education:		
Did not go to school, unknown	0.17	0.16
Primary school	0.73	0.73
Past primary	0.10	0.11
Father's occupational status		
Professional/administrative	0.14	0.17
Workers & agriculture workers	0.41	0.49
Farmers	0.27	0.19
Farmers (10+ ha)	0.08	0.05
Employer/self-employed	0.08	0.08
Other, unknown	0.01	0.02
Number of persons per heated room		

up to 1 person	0.05	0.19
1-2 persons	0.37	0.49
2-3 persons	0.28	0.22
3 and more persons	0.29	0.08
Unknown	0.01	0.02
Own SEP (at age 35)		
Siblings' Income		
1st decile	0.10	0.10
2nd decile	0.08	0.08
3rd decile	0.08	0.07
4th decile	0.08	0.07
5th decile	0.08	0.07
6th decile	0.08	0.08
7th decile	0.08	0.08
8th decile	0.08	0.09
9th decile	0.08	0.09
10th decile	0.09	0.11
Unknown	0.16	0.16
Siblings' education		
Basic	0.53	0.49
Upper secondary (lower track)	0.24	0.23
Upper secondary (higher track)	0.11	0.14

Lower-degree tertiary	0.04	0.05
Highest-degree tertiary	0.07	0.10
Siblings' home ownership		
No owner	0.30	0.28
Owns house/share	0.56	0.59
Unknown	0.14	0.14
Siblings' occupation		
Self-employed	0.11	0.11
Upper white-collar	0.12	0.15
Lower white-collar	0.26	0.27
Blue-collar	0.31	0.27
Other/unknown	0.20	0.19
Siblings' employment status		
Employed	0.71	0.72
Unemployed	0.02	0.01
Homemakers	0.06	0.06
Others/Unknown		

Appendix - Figures

Fig. A1 – Comparison of parametric and non-parametric estimate of the cumulative hazard function for all-cause mortality

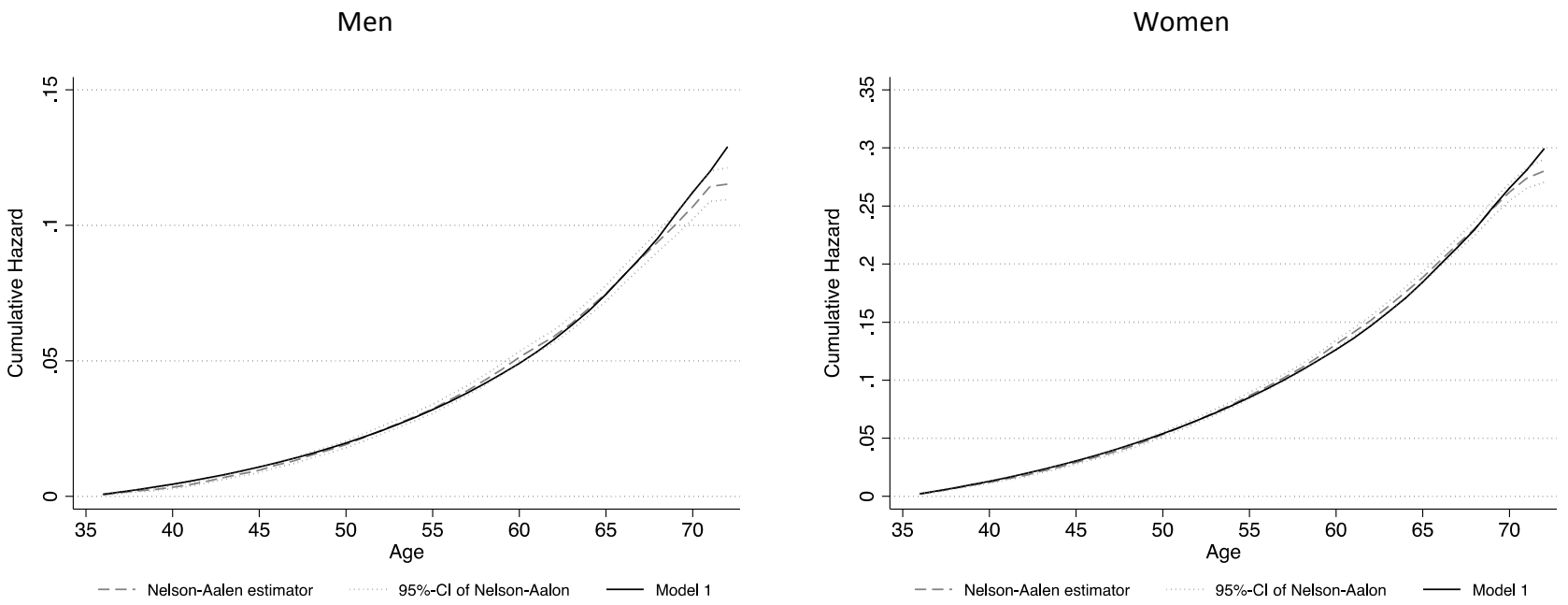


Fig. A2 – Equivalent years of ageing for men and women in the baseline model

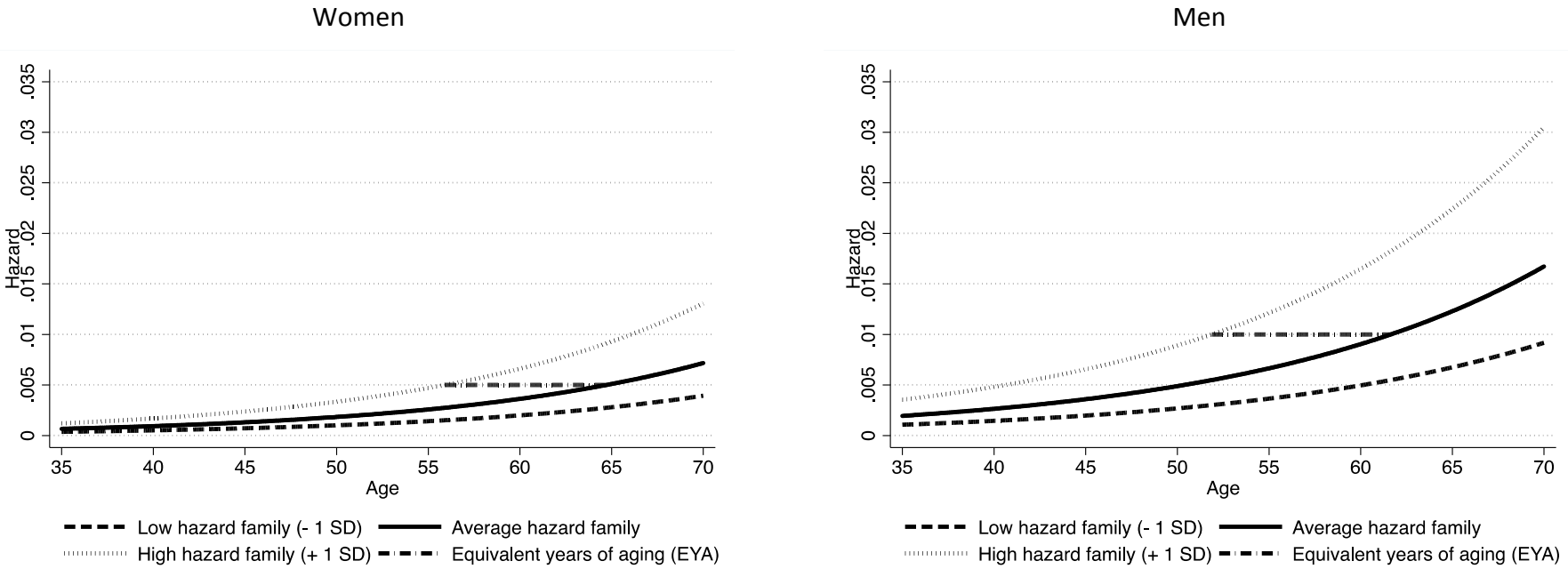


Fig. A3 – Distribution of EYA in the population based on model prediction for all-cause mortality

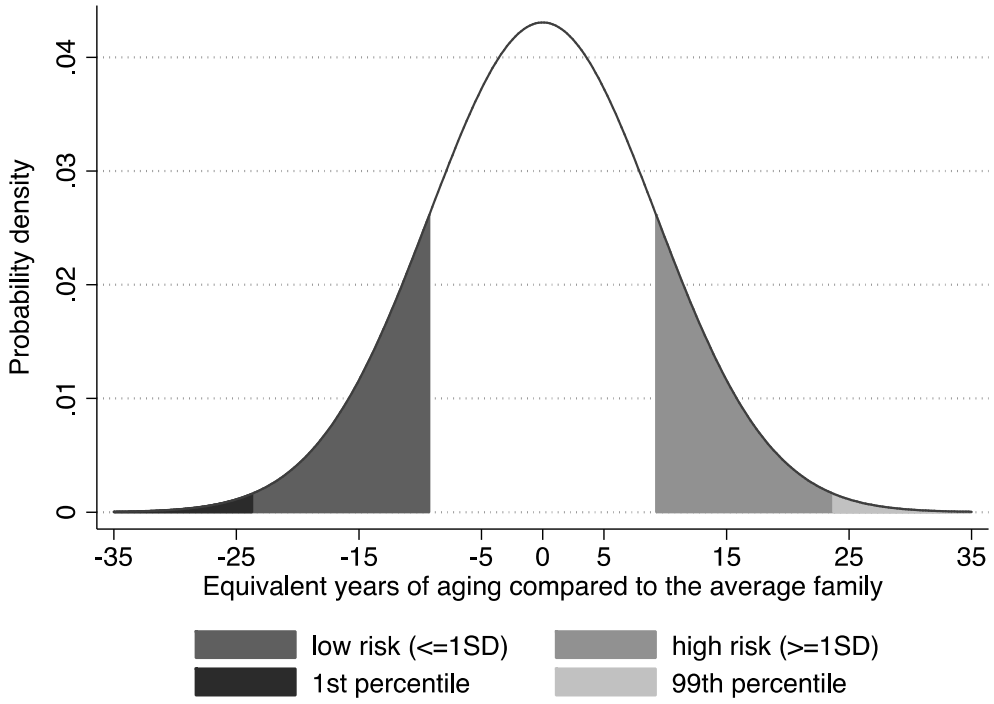


Fig. A4 – Differences in sibling similarity between models and by cause of death

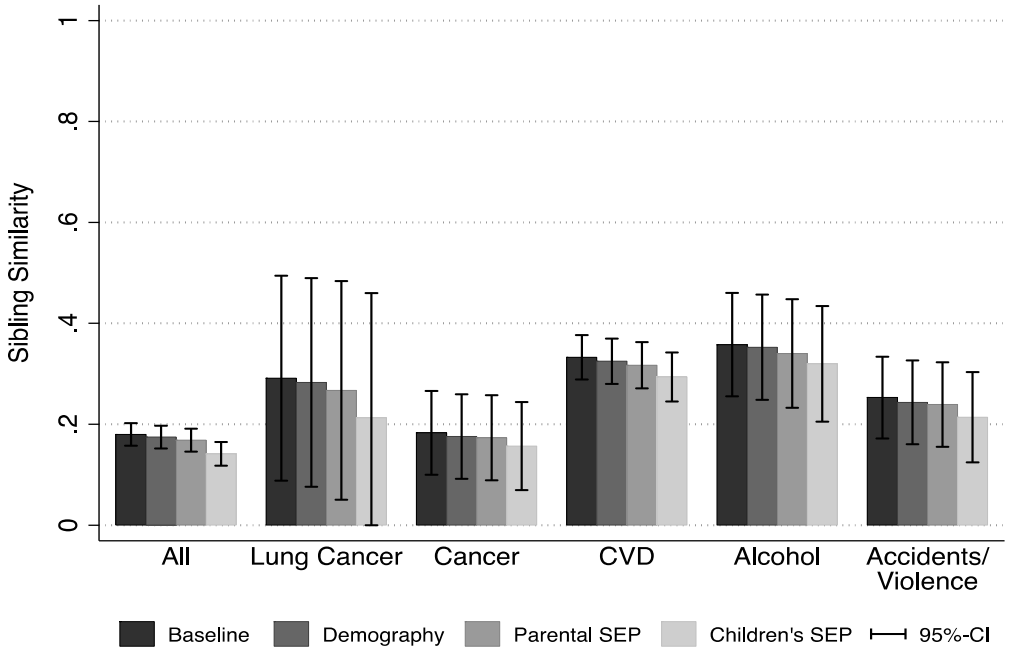
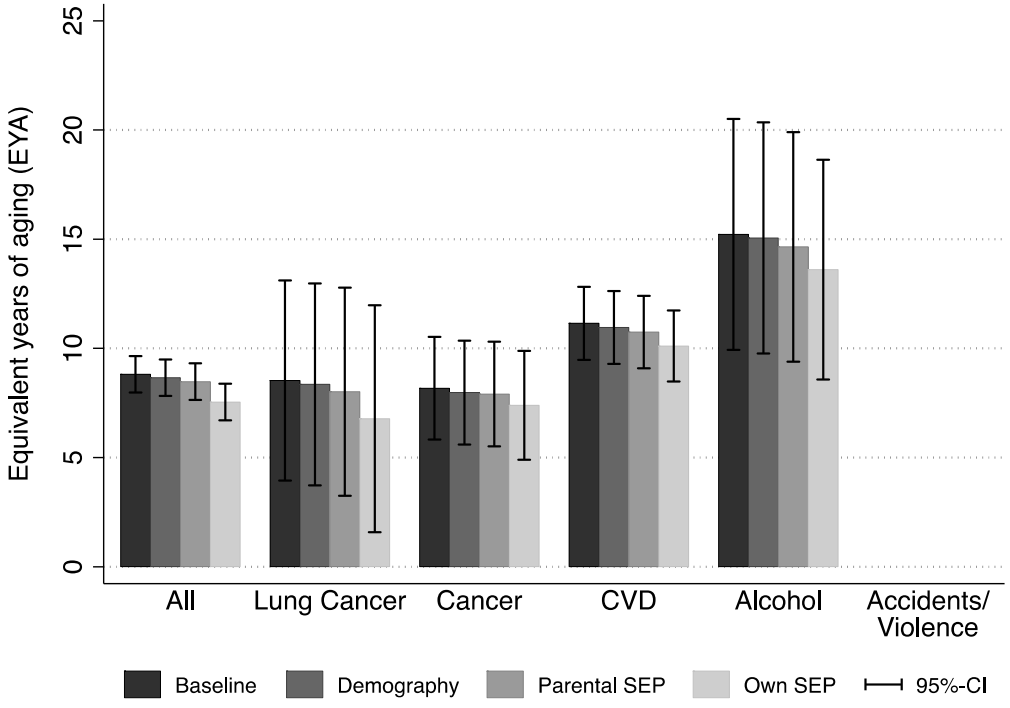


Fig. A5 – Differences in equivalent years of aging between models and by cause of death



Note: Estimates for accidents and violence are missing because they were unreliably high – the estimate for the shape parameter (the nominator) approached zero.

Fig. A6 – Sibling similarity for different models and distributional assumptions

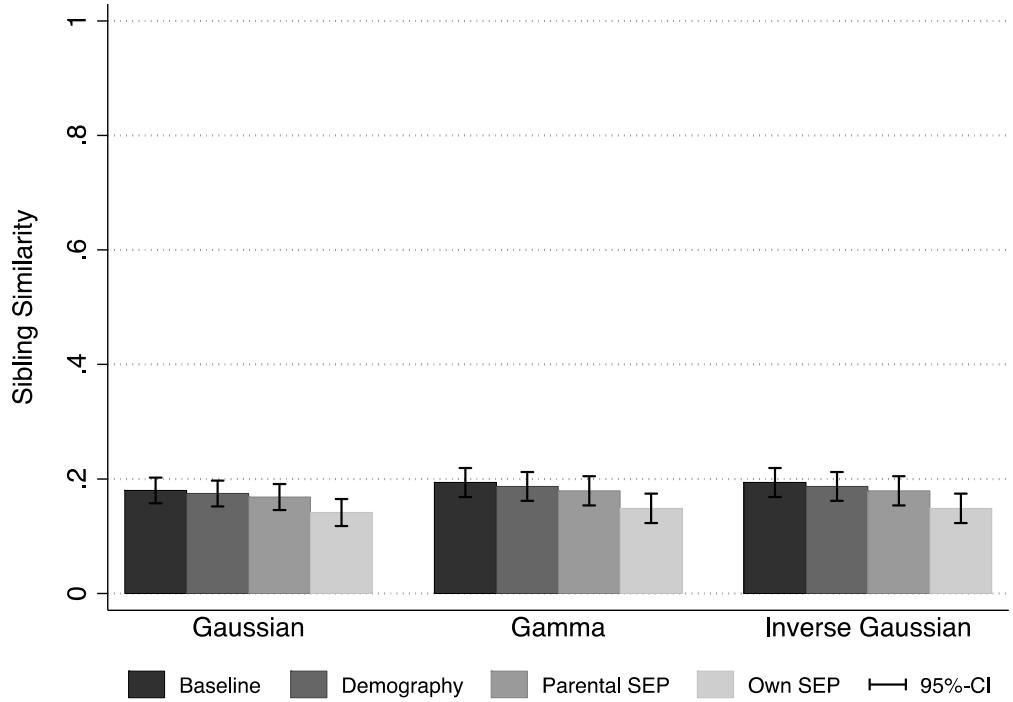


Fig. A7 – Equivalent years of aging for different models and distributional assumptions

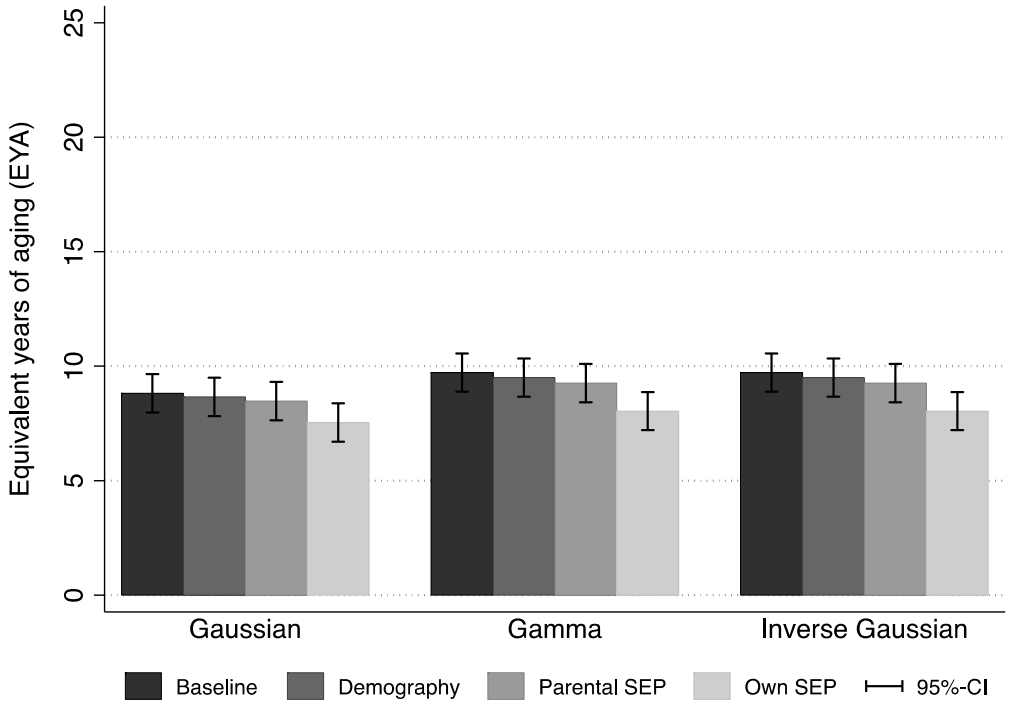
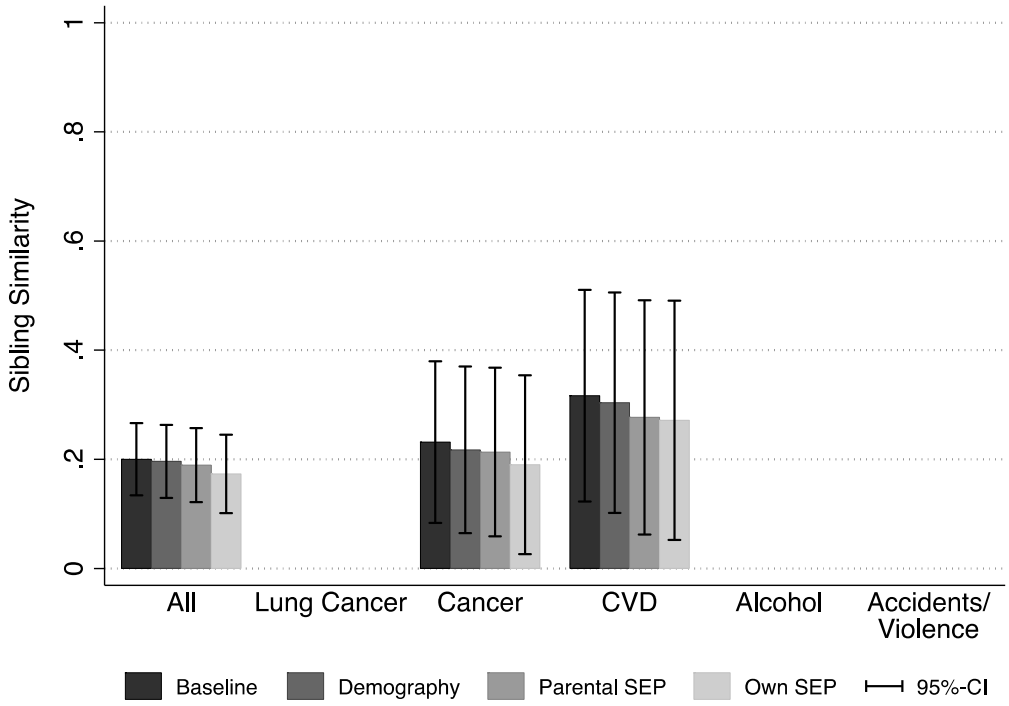
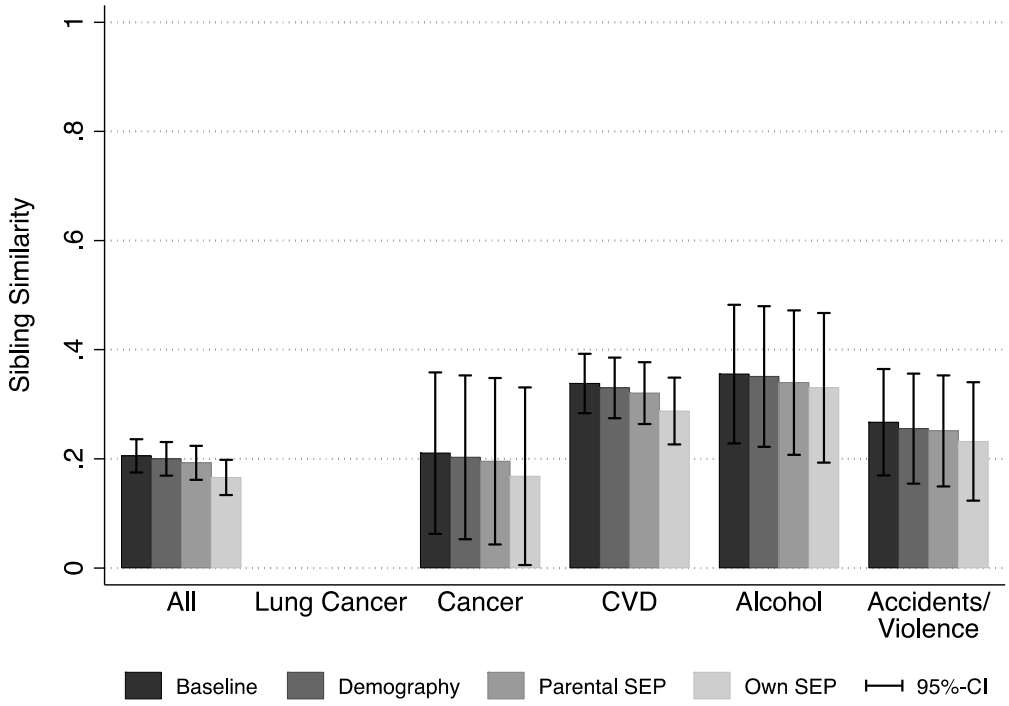


Fig. A8 – Sister similarity in all-cause and cause specific mortality



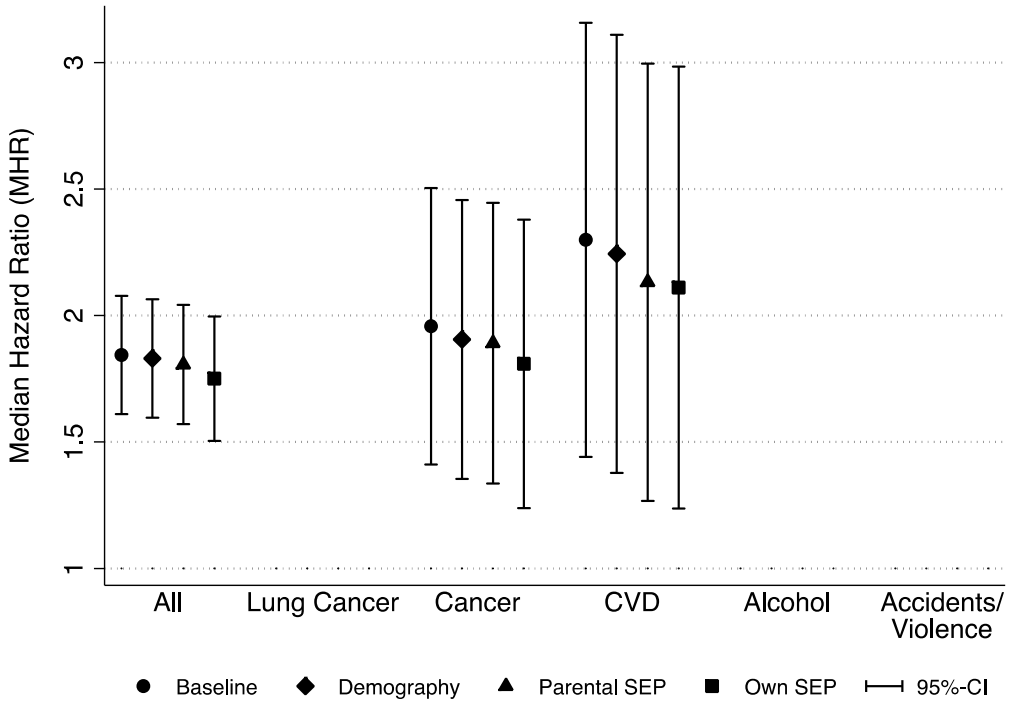
Note: There were not enough deaths related to lung cancer (<500) to get reliable estimates of the inter-family variation.

Fig. A9 – Brother similarity in all-cause and cause specific mortality



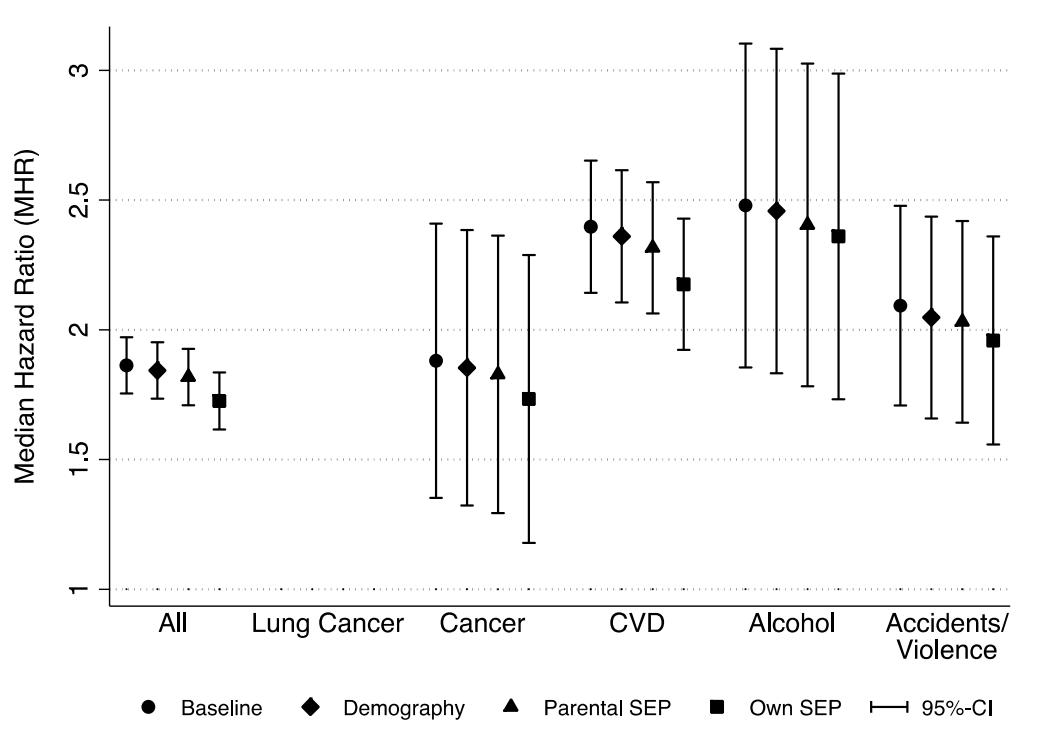
Note: There were not enough deaths related to lung cancer, alcohol, and accidents and violence (<500) to get reliable estimates of the between family variation.

Fig. A10 – Median hazard ratio in all-cause and cause specific mortality – Sisters



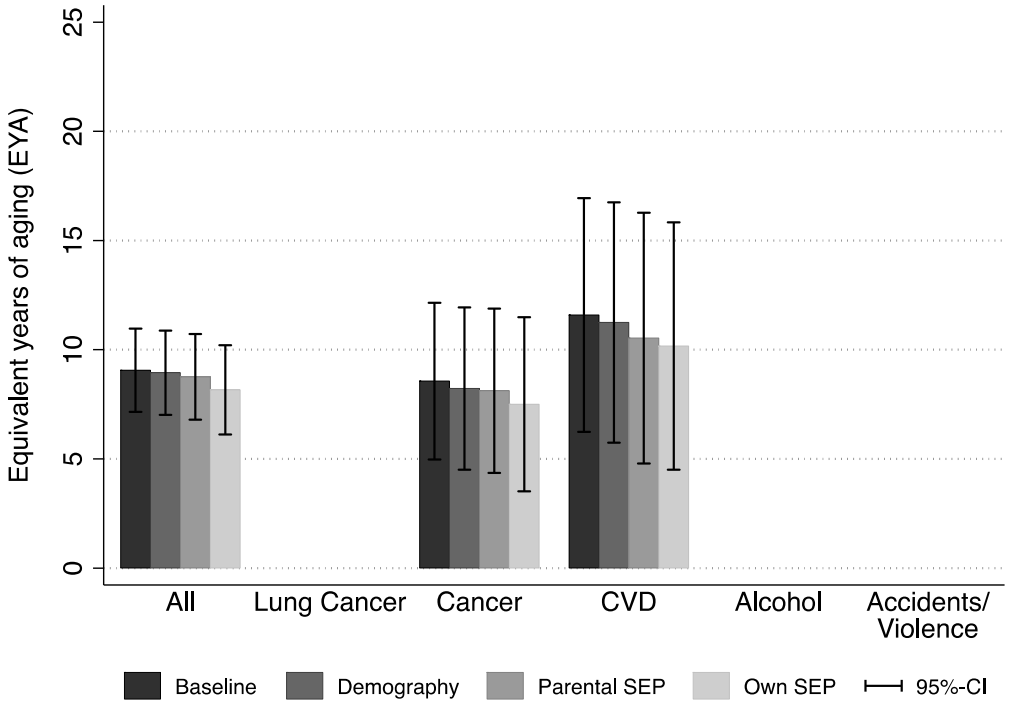
Note: There were not enough deaths related to lung cancer, alcohol, and accidents and violence (<500) to get reliable estimates of the inter-family variation.

Fig. A11 – Median hazard ratio in all-cause and cause specific mortality – Brothers



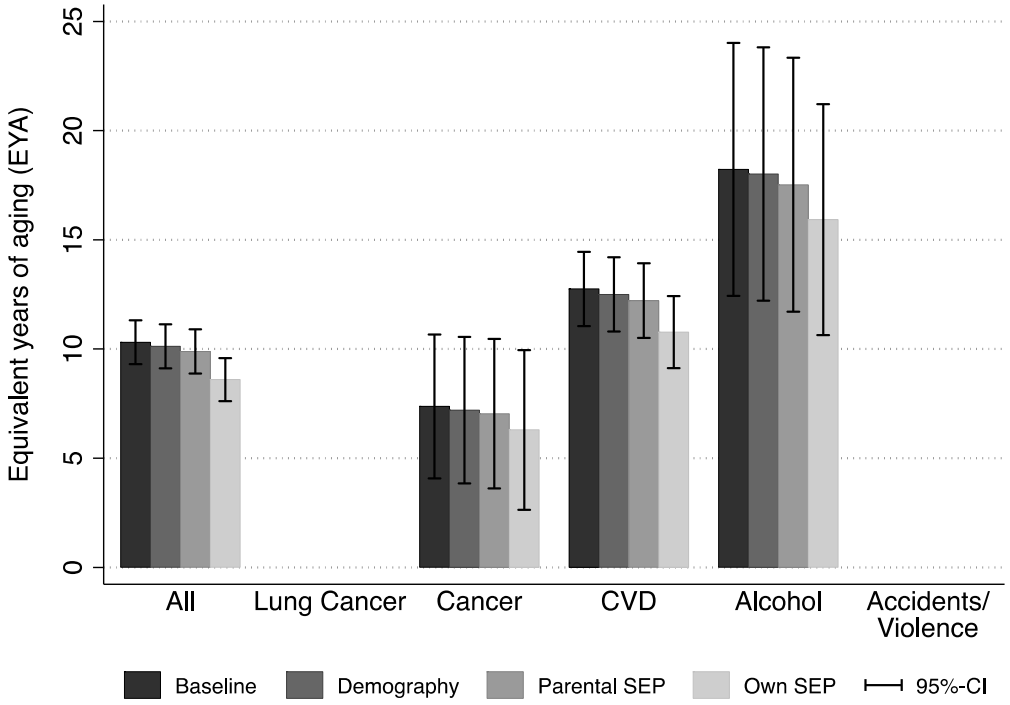
Note: There were not enough deaths related to lung cancer (<500) to get reliable estimates of the inter-family variation.

Fig. A12 – Equivalent years of aging in all-cause and cause specific mortality – Sisters



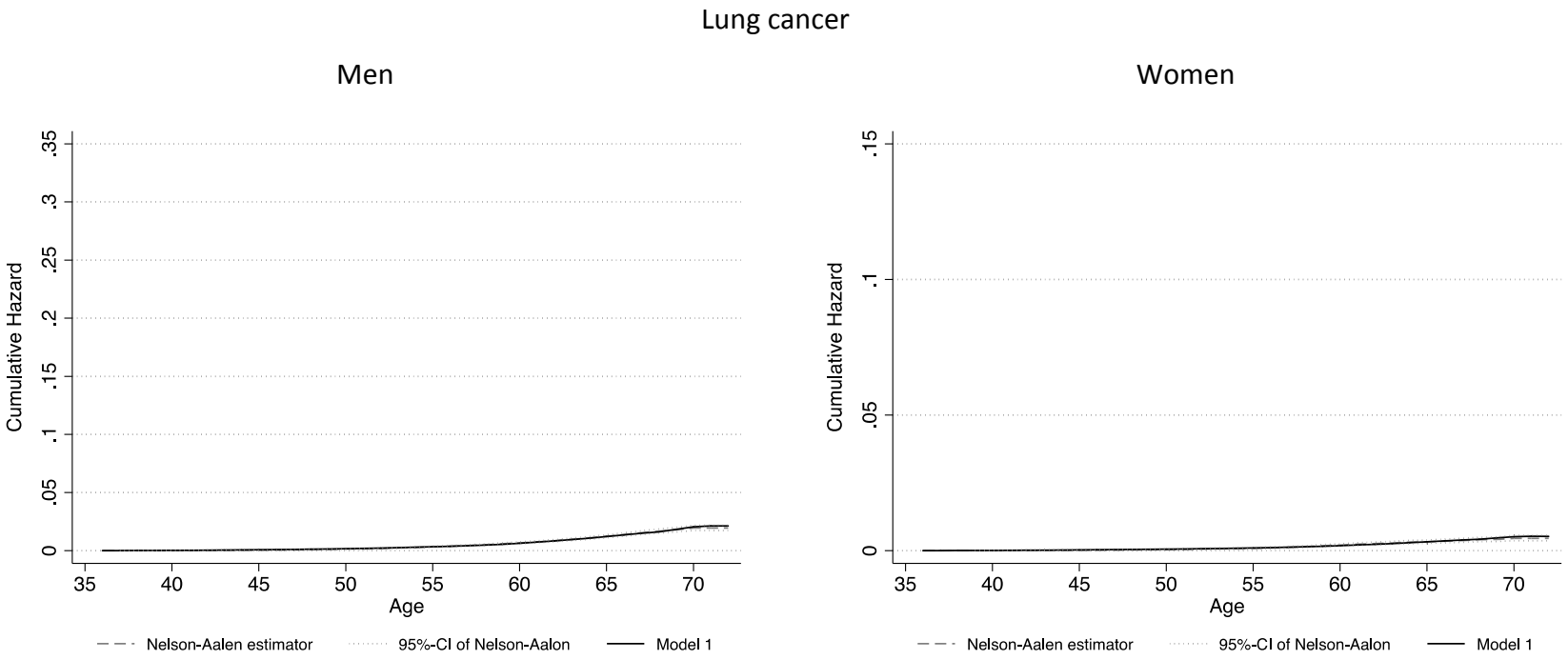
Note: There were not enough deaths related to lung cancer, alcohol, and accidents and violence (<500) to get reliable estimates of the inter-family variation.

Fig. A13 – Equivalent years of aging in all-cause and cause specific mortality – Brothers



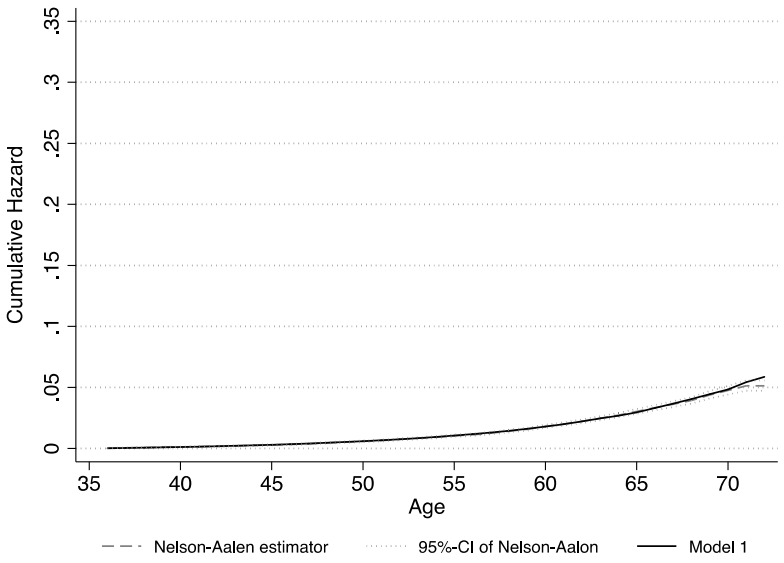
Note: There were not enough deaths related to lung cancer and accidents and violence (<500) to get reliable estimates of the between family variation.

Fig. A14 - Comparison of parametric and non-parametric estimate of the cumulative hazard function by causes of death

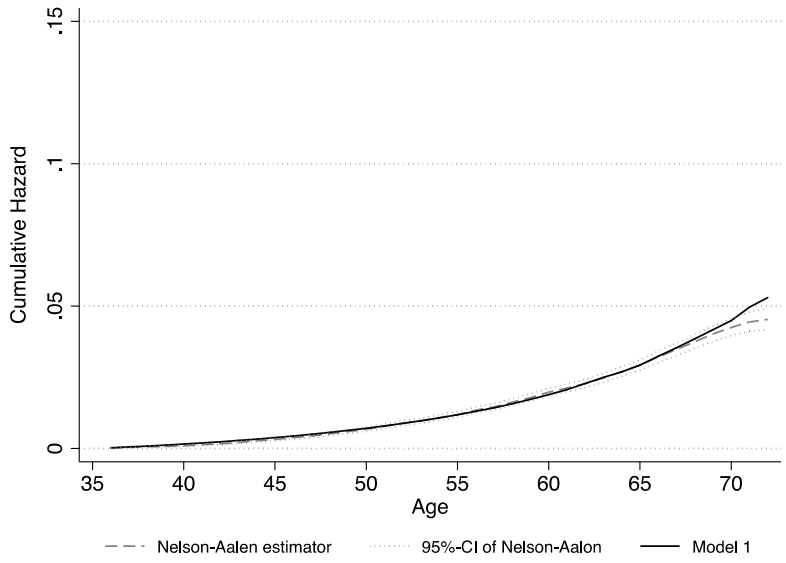


Cancer

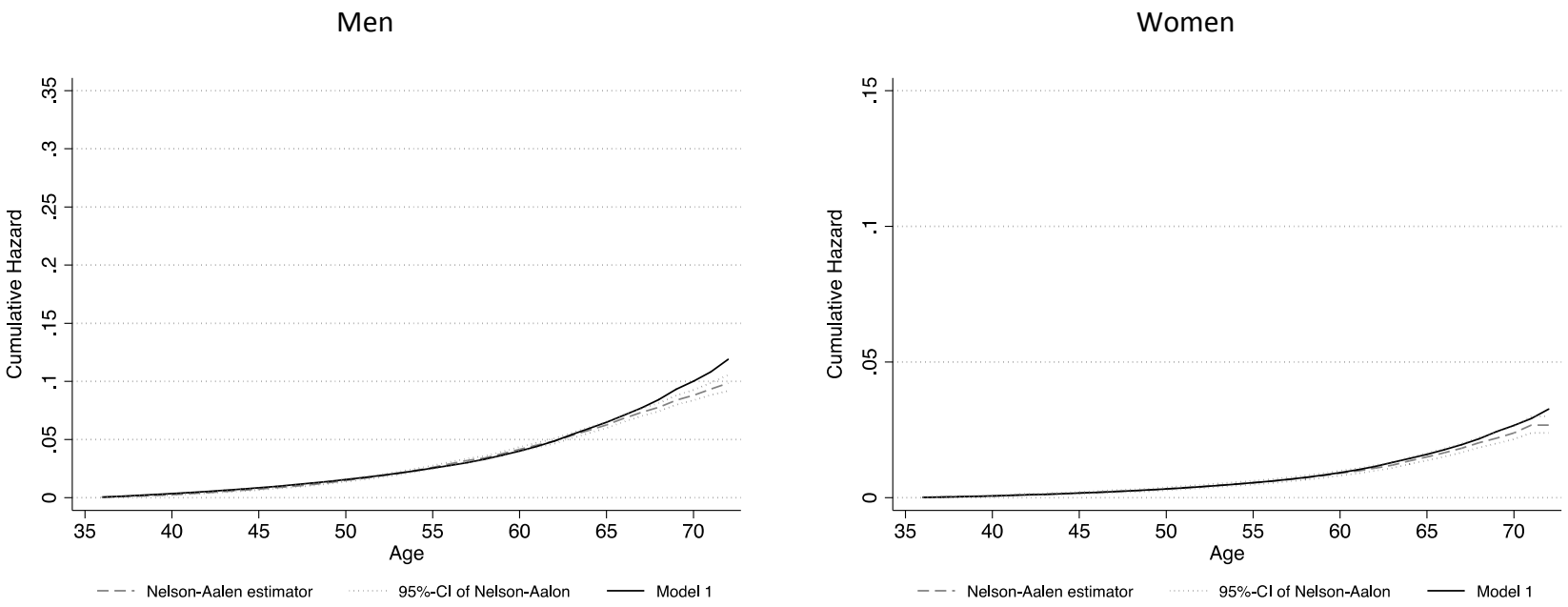
Men



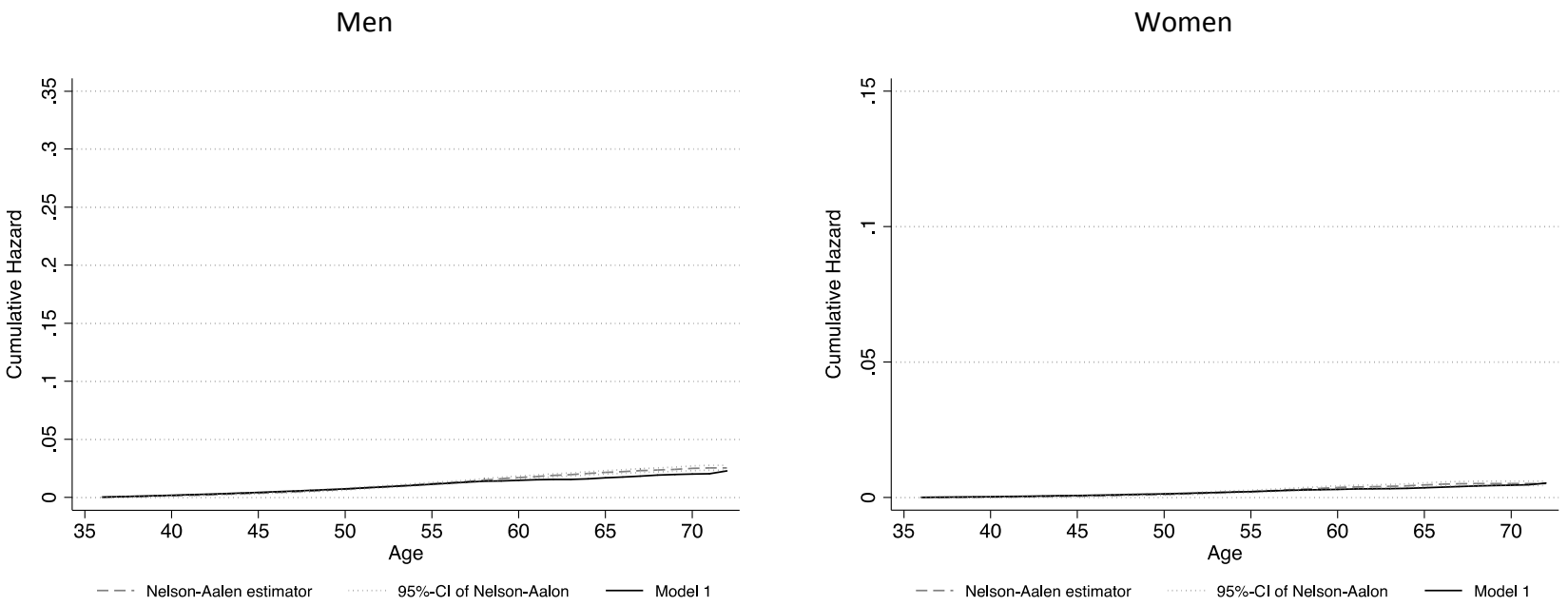
Women



CVD

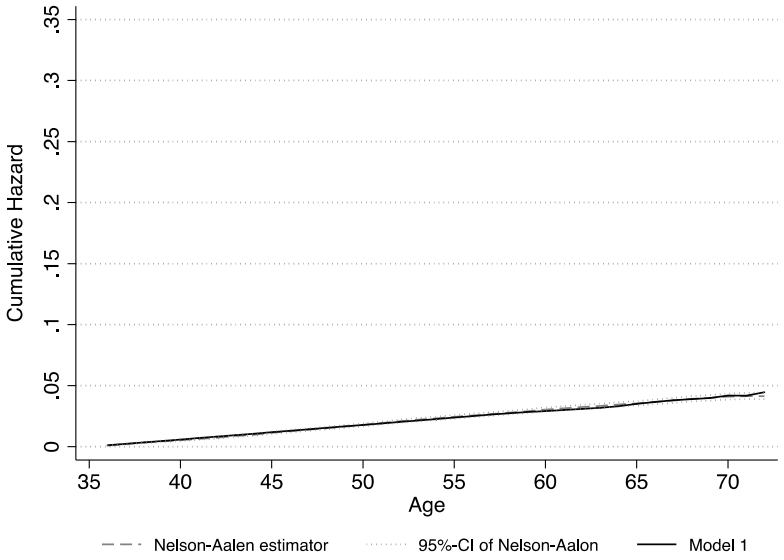

Men
Women
Cumulative Hazard
0
.05
.1
.15
.2
.25
.3
.35
35
40
45
50
55
60
65
70
Age
Nelson-Aalen estimator
95%-CI of Nelson-Aalon
Model 1

Alcohol


Men
Women
Cumulative Hazard
0
.05
.1
.15
.2
.25
.3
.35
35
40
45
50
55
60
65
70
Age
Nelson-Aalen estimator
95%-CI of Nelson-Aalon
Model 1

Accidents and violence

Men



Women