

DISCUSSION PAPER SERIES

IZA DP No. 14507

**Educational Differences in Mortality
and Hospitalisation for Cardiovascular
Diseases for Males**

Govert E. Bijwaard

JUNE 2021

DISCUSSION PAPER SERIES

IZA DP No. 14507

Educational Differences in Mortality and Hospitalisation for Cardiovascular Diseases for Males

Govert E. Bijwaard

*Netherlands Interdisciplinary Demographic Institute
and IZA*

JUNE 2021

Any opinions expressed in this paper are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but IZA takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The IZA Institute of Labor Economics is an independent economic research institute that conducts research in labor economics and offers evidence-based policy advice on labor market issues. Supported by the Deutsche Post Foundation, IZA runs the world's largest network of economists, whose research aims to provide answers to the global labor market challenges of our time. Our key objective is to build bridges between academic research, policymakers and society.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

Educational Differences in Mortality and Hospitalisation for Cardiovascular Diseases for Males*

High educated individuals are less frequently admitted to hospital for cardiovascular diseases and live longer than the lower educated. We address whether the educational gradient in the mortality rate can be explained by the educational difference in the timing of CVD hospitalisation. We account for possible selective hospitalisation, by using a correlated multistate hazard model (a 'Timing-of-events'- model) and, for selection into education, by using inverse propensity weighting based on the probability to attain higher education. We use Swedish Military Conscription Data (1951-1960), for males only, linked to administrative Swedish registers. Our empirical results indicate a clear educational gradient in mortality and in the impact of CVD hospitalisation on mortality. The implied educational gain in the number of months lost is, however, mainly due to other factors than CVD hospitalisation. Extending the analysis to cause specific mortality reveals that the largest educational differences exist in death due to external causes.

JEL Classification: C41, I14, I24

Keywords: education, mortality, CVD hospitalisation, timing-of-events, inverse propensity weighting

Corresponding author:

Govert E. Bijwaard
Netherlands Interdisciplinary Demographic Institute
University of Groningen
PO Box 11650
2502 AR The Hague
The Netherlands
E-mail: bijwaard@nidi.nl

* Acknowledgements. I thank Mikko Myrskylä, Per Tynelius and, attendants at the Nordic Demographic Symposium in Reykjavik for their helpful comments.

1 Introduction

The strong association between education and health is one of the most widely studied in economics. Observational evidence suggests that high educated people live longer (Mazumder, 2008, Clark and Royer, 2013, Fletcher, 2015, McCartney et al., 2013). The recent focus in the education-health literature is on determining to what extent education causes later mortality (Galama et al., 2018, Xue et al., 2021). However, uncovering the underlying mechanisms that produce these causal relationships are mostly ignored. One path to higher mortality for the low educated is through more (and longer) hospitalisations, in particular hospitalisations for cardiovascular diseases (CVD). CVD hospitalisation is clearly associated with life-style behaviour, as smoking, limited exercise and obesity are all well-known causes for cardiovascular health problems.

Higher education may have countervailing effects on hospitalisation if it reduces the probability of negative health conditions that might lead to hospitalisation yet increases the probability of hospitalisation (e.g., through greater income, more knowledge, or better connections) for given health conditions. Likewise, higher education may have countervailing effects on mortality if it increases income and access to health-system care but also increases higher-risk behaviors and selection into more stressful occupations.

Only a few studies have attempted to identify the causal effect of education on hospitalisation (Arendt, 2008, Tansel and Keskin, 2017, Meghir et al., 2018). Arendt (2008) found, for Denmark, using a bivariate probit model, a clear educational gradient in the probability of ever been hospitalised, but no significant educational impact on the number of days in hospital. Meghir et al. (2018) found, for Sweden using linear Difference-in-difference regressions and using a regression discontinuity approach both based on a reform in compulsory schooling, no impact of education on the total number of days in hospital nor on the probability of ever been hospitalised for (amongst other diseases) circulatory diseases. Tansel and Keskin (2017) found, for Turkey using a Tobit and a Double Hurdle model, that an increase in years of education reduces the number of days hospitalized. However, in all these papers the role hospitalisation experience plays in explaining mortality is ignored.

The contribution of this paper is threefold. First, it addresses the educational gradient in the mortality rate and the impact of the timing of entry and discharge of the life-style related

CVD hospitalisation on the educational gradient. Second, it derives the implied educational gain, both directly and through CVD hospitalisation, on survival and months lost, of improving education with one level. Third, it is the first paper that investigates the hospitalisation and mortality process jointly, through a correlated multistate hazard model.

Most studies investigating the impact of education on mortality have used a linear or a probit model to estimate the educational gradient, although the age at death (and timing of CVD hospitalisation process) is clearly a duration outcome. In duration analysis the hazard rate, the instantaneous probability that an individual enters a certain state (death, hospital-entry or discharge) at a certain age conditional on surviving (or not in that state) up to that age, is usually modelled. Accounting for right-censoring, when the individual is only known to have survived up to the end of the observation window and time-varying variables, e.g. the age somebody enters hospital, are easy to handle in hazard models, see a.o. Lancaster (1990) and Van den Berg (2001). A common way to accommodate the presence of observed characteristics is to specify a proportional hazard (PH) model, in which the hazard is the product of the baseline hazard, the age dependence, and a log-linear function of covariates. Neglecting confounding in inherently non-linear models, such as proportional hazard models, leads to biased inference. To accommodate this (see e.g. Van den Berg (2001) for a discussion of the importance of this), the mixed proportional hazard model (MPH) extends the PH model by multiplying it by a time-invariant person-specific random error term. This has been the main model for analysis of duration data in economics.

Only a few studies on the educational gradient in mortality used a hazard model. Meghir et al. (2018) used a regression discontinuity-type approach based on a reform in compulsory schooling, Bijwaard et al. (2017) and Bijwaard and Jones (2019) used an inverse propensity weighting method and, Bijwaard et al. (2015a,b, 2019) a structural modelling approach. Yet, none of these studies considered the relation between hospitalisation and mortality or accounted for selective hospitalisation (nor the timing of hospitalisation) in the mortality process. Bijwaard and van Kippersluis (2016) has shown that education influences both entry and discharge from hospitals, and that higher educated individuals are less likely to die after a hospitalization. When accounting for the role of intelligence using a structural equation multistate model, this association disappears. In their model the interdependence of the hospitalisation and mortality process, is solely driven by intelligence. They ignore the

endogeneity of the timing of hospitalisation.

The richness of our data enables us to go beyond standard modelling of life cycle durations, and to tackle the complex task of examining jointly the mortality and hospitalisation processes. Correlation between these two processes might stem from correlated unobserved heterogeneity, and our model accommodates this. Applying the “timing-of-events” (ToE) method (Abbring and van den Berg, 2003) allows us to give the estimated effects of the hospitalisation dynamics on mortality a causal interpretation. Controlling for unobserved correlated heterogeneity in the hospitalisation and mortality processes is thus crucial, since otherwise the resulting endogeneity would confound the causal impact. In particular, (first-time) hospital entry, hospital discharge, hospital re-entry and mortality durations are modelled as mixed proportional hazards which incorporate correlated unobservables.

However, the timing-of-events method still fails to correct for possible endogeneity of educational attainment. The propensity score method (Hirano et al., 2003, Caliendo and Kopeinig, 2008) we employ accounts for such confounding. It is based on the unconfoundedness assumption, which assumes that all variables that affect mortality, hospitalisation and education attainment are observed. This is a stringent assumption, but our data contain important factors as detailed family socioeconomic background (including paternal- and maternal socioeconomic status at birth and education level), cognitive skills (IQ-test) and non-cognitive skills (psychological test). However, confounding between the education choice and the hospitalisation-mortality process may still exist. To test possible violation of the unconfoundedness we use an extension of the sensitivity approach of Bijwaard and Jones (2019).

We base our (generalised) propensity scores on a multistage sequential model of educational attainment developed by Cameron and Heckman (2001) with five education levels. Based on the estimated propensity scores we estimate a weighted Timing-of-events model (IPW-ToE) with weights based on the inverse of the generalised propensity score (Inverse Propensity Weighting, IPW, (Hirano et al., 2003)), see Bijwaard and Jones (2019) for an application of (a single valued) IPW method for a Mixed Proportional Hazard mortality model. Based on the estimated model we calculate the educational gain of improving education in both the survival probability till age 63 (the highest age observed) and in the number of months lost till age 63. We also decompose these educational gains into an indirect effect, running through changes in the hospitalisation process, and a direct effect due to other factors.

Data from the Swedish Military Conscription Data (1951-1960), linked to administrative Swedish registers, offers the opportunity to investigate the impact of hospitalisation and education on (cause-specific) mortality. We have information on about half a million men who are followed from the date of conscription till the end of 2012, or till death. For those men who die we observe the cause of death. From the Swedish National Hospital Discharge Register we observe CVD hospital care from 1964 till the end of 2012. These data include recordings of demographic and socioeconomic characteristics such as education, parental socioeconomic status, parental education, along with anthropometric measures, an intelligence test and a psychological assessment. Educational level was classified in five categories: primary education; some secondary education (2 years); full secondary education (3 years); post-secondary education and higher education.

A couple of other studies have also used data from Sweden to investigate the educational-mortality gradient. Spasojević (2010), Lager and Torssander (2012) and Meghir et al. (2018) all used the compulsory schooling reform 1949–1962 in Sweden as an instrument for an increase in education, Spasojević (2010) using survey data and both Lager and Torssander (2012) and Meghir et al. (2018) using register data of children born 1940–1955. Lundborg et al. (2016) used a (linear) twin fixed effects model design, twins born 1886–1958, to address endogeneity of education choice. Bijwaard et al. (2017) and Bijwaard et al. (2019) both focus on the educational gradient in cause-specific mortality. Bijwaard et al. (2017) used a cause-specific months-lost model with IPW and family fixed effects and Bijwaard et al. (2019) used a structural cause-specific hazard model to account for educational endogeneity. The data used in these papers is close to the data we used, the Military Swedish Conscription data linked to the death register and census data. Bijwaard et al. (2017) used the birth cohorts 1951–1983, but restricted the analyses to men with at least one brother, while Bijwaard et al. (2019) used, just as we do, only the birth cohorts 1951–1960 but discarded the men without IQ measurement. However, neither of these papers used information on CVD hospitalisation.

The empirical analyses show that both education and CVD hospitalisation are important factors influencing mortality. We find a clear educational gradient in the mortality, even after accounting for the endogeneity of education through inverse propensity weighting. We also find that mortality is much higher for those in hospital for CVD (39-68 times higher) and for those who have been in hospital (2.1 to 4.7 times higher). These hospitalisation effects

also exhibit a educational gradient, with a decreasing impact of hospitalisation the higher the education level.

Based on these parameter estimates we calculate the educational gains and find that men with only primary education would gain the most (about 9 months from age 18 to 63) if they had a higher education level. Only a small part of this educational gain is running through a change in CVD hospitalisation. Although men with post-secondary education would only gain 2.3 months if they had higher education most of this educational gain (1.6 months) is attributable to changes in the CVD hospitalisation for the higher educated.

Evidence suggests differential impact of education on various diseases accumulating in different educational cause-specific mortality gradients (Galobardes et al., 2004, Bijwaard et al., 2017, 2019). As the socioeconomic association seems the largest for cardiovascular diseases most studies have focussed on socioeconomic differences in mortality on these types of diseases. Some have indeed found that the incidence of cardiovascular disease is higher for individuals with low socioeconomic status (Mackenbach et al., 2008, Kulhánová et al., 2014). However, Bijwaard et al. (2017, 2019) found that most of the educational gains in mortality up till age 63 (the same maximum age we are using) are attributable to the reduction in mortality due to external causes and that the reduction in death due to CVDs with improving education is rather small.

To investigate whether education and hospitalisation for CVD affects different causes of death differently we also estimate a model with cause-specific mortality rates, distinguishing five different causes of death : 1) Ischemic Heart Disease (IHD); (2) Stroke; (3) other cardiovascular causes; (4) External causes, and (5) Other (natural) causes of death. The model is an extension of the timing-of-events model with IPW. We find that all causes of death show a clear educational gradient (except between post-secondary and university/PhD), with the largest educational differences for external causes of death. We also find that death due to other CVD is affected the most by hospitalisation. However, hospitalisation for CVD also affects the mortality due to other natural causes and due to external causes.

2 Data

The data come from several Swedish population-wide registers which are linked using unique individual identification. The Swedish Military Conscription Data includes demographic information of the conscripts and information obtained at the military examination, including a battery of intelligence tests and a psychological assessment. These data are linked to the National Population and Housing Censuses (1960-1990), containing information on the socioeconomic status and educational levels of the parents. Information on each conscript's own education was obtained from the Longitudinal Integration Database for Health Insurance and Labour Market Studies (LISA) for the 1990-2012 period, and information on cause-specific mortality as the underlying cause of death was obtained from the Cause of Death Register for the period up to 2012. The information (timing of admittance and discharge) on hospitalisation for Cardiovascular Diseases (CVD)¹ is derived from the Inpatient Register. Coverage runs from 1964 to December 31st 2012. The study population consists of men born between 1950 and 1960, who were identified in the Multi-Generation Register, and who were conscripted for military examination between 1969 and 2001 usually when they were aged 18-20. At that time in Sweden, military service was mandatory for men only. So, we only observe males. We selected only those, 517,843, men for whom at least one parent is known and for whom we observe the conscription date.

These data include recordings of demographic and socioeconomic characteristics such as education, parental socioeconomic status, parental education, along with anthropometric measures, an intelligence test and a psychological assessment and health measures (height, weight, blood pressure, and muscular strength). The intelligence measurement is based on a battery of IQ tests, which consisted of four subtests that measured logical, spatial, verbal, and technical abilities. Each subtest was first evaluated on a normalized ninepoint (stanine) scale. The subtest scores were summed to obtain an overall score and transformed onto a stanine scale with a mean of five and standard deviation of two. We only used this final global IQ measurement. The psychological assessment is also based on a normalized ninepoint (stanine) scale. For both the intelligence and the psychological assessment measurements we also define a missing indicator, when this measurement is not observed. The parental socio-economic status has

¹ICD 8 and 9: 390-459 ICD 10: I.

seven categories: Non-manual workers at higher level, Non-manual workers at intermediate level, Non-manual workers at lower level, Farmers, Skilled workers, Unskilled workers and Others, not classified. We also define the unknown SES indicator. Parental education has six categories: primary (< 9 years), primary (9 – 10 years), Secondary education (2 years), Full secondary education (3 years), Post-secondary (University < 3 years) and Higher (University ≥ 3 years or PhD).².

We aggregated the observed education of the conscripts into five classes: (i) Less than 10 years of education (only primary schooling); (ii) Some secondary education (2 years); (iii) Full secondary education (3 years); (iv) Post-secondary education (less than 3 years) and (v) Higher education (University and PhD). More detailed information on the data can be found in Bijwaard et al. (2017).

About 80 thousand (of the 518 thousand) men have experienced CVD hospitalisation. Table 1 presents the distribution of this hospitalisation by educational attainment. The CVD hospitalisation experience is lower for the high educated and so is the average number of days spent in hospital for CVD. For men without hospitalisation experience low educated men have died four times more often than high educated men. While for men with hospitalisation experience the educational gradient is less steep. Finally, the table shows that hospitalisation experience clearly increases mortality.

Table 1: Descriptive statistics on hospitalisation and mortality ($N = 517,843$)

	Primary	Secondary education some	Post-secondary full	Higher (< 3 years)	Higher
% with hospitalisation	18.2%	16.4%	14.9%	13.6%	11.7%
Av.# of days in hospital	7.3	5.9	4.9	4.0	3.1
% died without hospitalisation	9.1%	5.0%	3.7%	2.5%	2.2%
% died with hospitalisation	14.4%	11.4%	9.5%	7.0%	6.1%

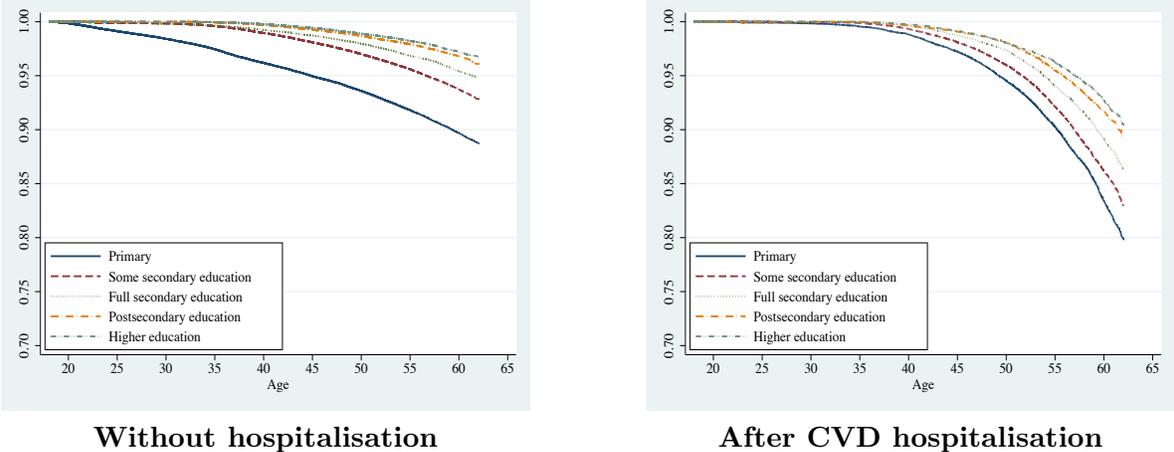
Next, we calculate the Kaplan-Meier survival curves. These non-parametric survival curves for the five education categories are shown in Figure 1 and reflect the mortality differences by education and by hospitalisation experience. Survival increases with the education level and the differences between the education levels increase with age. Comparing the survival curves without hospitalisation (left panel) and with hospitalisation (right panel) shows the

²Full demographic and childhood family characteristics at the time of military examinations by education level are presented in Table B.1, for men without hospitalisation for CVD and in Table B.2, for men with hospitalisation for CVD in Appendix B

large impact of hospitalisation on mortality. The Kaplan-Meier survival curves for the first admittance to hospital and re-admittance to hospital by education level in given in Figure 2 also show a clear educational gradient.

However, it takes time to experience hospitalisation, which depends on educational attainment and, these mortality differences, therefore, do not necessarily reflect the influence of education and/or hospitalisation on mortality per se. The observed differences in survival between low and high educated men could also be induced by a higher IQ or a higher socio-economic background of the high educated men. For example, understanding a doctor’s advice and adhering to complex treatments after hospitalisation may be driven by intelligence rather than education. In the next section we explain how we account for this.

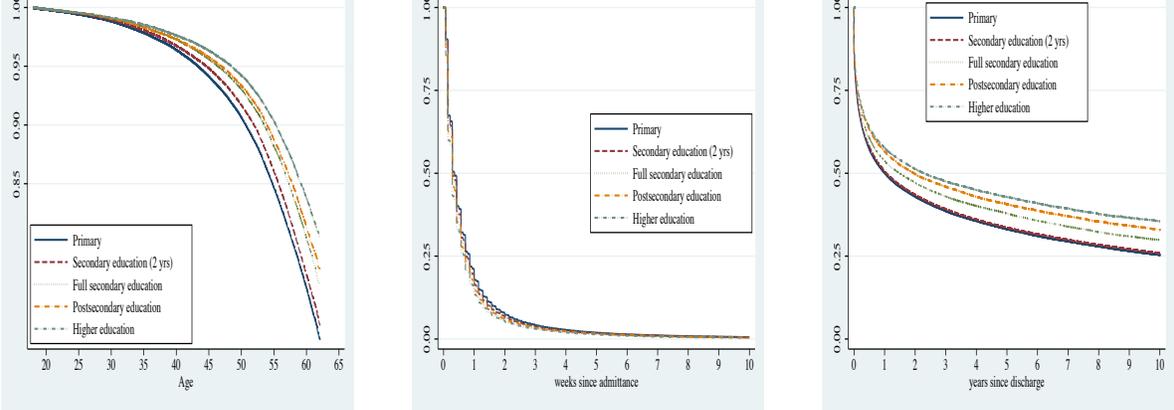
Figure 1: Kaplan-Meier mortality survival curves by education level



Without hospitalisation

After CVD hospitalisation

Figure 2: Kaplan-Meier hospitalisation survival curves by education level



first hospitalisation

discharge

re-admission

3 Method

Dynamic selection and censoring are inherent to a duration outcome, such as age at death and the age at (CVD) hospitalisation, and need to be accounted for in the modelling. Our main outcome measure is the mortality hazard rate, the instantaneous probability that an individual dies at a certain age conditional on surviving up to that age. Accounting for right-censoring, when the individual is only known to have survived up till the end of the observation window, and left-truncation, when only those individuals are observed who were alive at a certain time (military examination), are easy to handle in hazard models (Van den Berg, 2001). A common way to accommodate the presence of observed characteristics is to specify a proportional hazard (PH) model, in which the hazard is the product of the baseline hazard, the age dependence, and a log-linear function of covariates. In a PH model it is also easy to incorporate time-varying variables, like hospitalisation and hospitalisation experience. However, it is hardly ever possible to include all relevant factors, either because the researcher does not know all the relevant factors or because it is not possible to measure them. Ignoring such unobserved heterogeneity or confounding may have a huge impact on inference in proportional hazard models, see e.g. Van den Berg (2001). A common solution is to use a Mixed Proportional Hazard (MPH) model, in which it is assumed that all unmeasured factors and measurement error can be captured in a multiplicative random term.

A major methodological concern with the empirical analysis of the impact of hospitalisation on mortality is that the admittance and discharge processes depend on individual characteristics, both observed and unobserved, that also influence mortality. This implies that any observed relationship between admittance to (or discharge from) hospital and mortality may be caused by unobserved factors that influence both the hospitalisation and mortality. For example, a finding that men with high educated fathers live longer may not necessarily imply that low socioeconomic background (low educated father) causes higher mortality. Rather, it may be induced by the higher hospitalisation of men from low socio-economic background. To account for the interdependence of the hospitalisation process we model the first admittance-, discharge- and re-admittance hazards of this process simultaneously with the mortality hazard. This is a multistate model with correlated hazards, also called a ‘timing-of-events model’ (ToE) (Abbring and van den Berg, 2003), which explicitly controls for the correlation between

the hospitalisation process and mortality, to account for this interdependence.

3.1 Timing-of-events method

Let T^m denote the time, age, till death (mortality), T^h the age at the start of the first hospital spell, T^d the age at hospital discharge and, T^r the age at hospital re-admittance (after discharge). The durations of a hospital stay and the time after hospitalisation are denoted by $d^h = T^d - T^h$ and $d^r = T^r - T^d$. In order to keep track of hospitalisation events, we also define the associated time-varying indicators: the indicator $I^h(t)$ takes value one if the individual is in hospital at age t , and $I^o(t)$ indicates that the individual has experienced a period of hospitalisation before age t .

In Figure 3(a) we depict the hospitalisation dynamics and mortality for an arbitrary military recruit. This recruit is not in hospital (healthy) at the military examination, as holds for all observed men. At some age t^h the recruit is admitted to hospital for a CVD. This implies that the time till first hospitalisation is $\delta^h = t^h$. The recruit is discharged from hospital at age t^d . Thus, the time he stayed in hospital is $\delta^d = t^d - t^h$. However, at age t^r he is admitted again to hospital for a CVD, which implies the time till re-hospitalisation was $\delta^r = t^r - t^d$. At age t^m this recruit dies, after a time of $\delta^d = t^m - t^r$ in hospital. During hospitalisation the indicator I^h is one and after the first hospitalisation spell indicator I^o is one. Thus, both I^h and I^o are one when the recruit is in hospital for the second time.

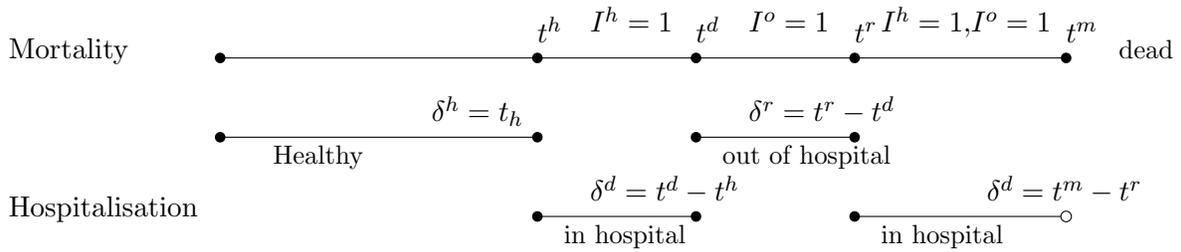
The most commonly applied multistate model in biostatistics is the illness-death model (Hougaard, 2000, Putter et al., 2007). In this type of model individuals start out healthy. From healthy they may become ill (enter hospital) or they may die. Ill individuals may die or recover (leave hospital) and become healthy again. The timing-of-events model is a special case of a such a model, in which the mortality not only depends on current illness-status, but also on illness experience. This model is depicted in Figure 3(b). Another large difference with standard ill-death models is that in the Timing-of-Events model the transition rates among the different states are independent conditional on unobserved random heterogeneity.

We model the first admittance to hospital using a Mixed Proportional Hazard (MPH) model

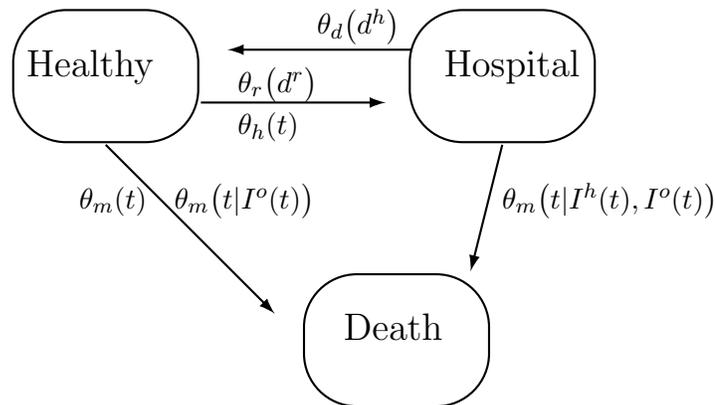
$$\theta_h(t|x, v) = v_h \lambda_{h0}(t) \exp\left(x\beta_x^h + e\beta_e^h\right) \quad (1)$$

Figure 3: The Timing-of-events (ToE) Model : how the timing of the hospitalisation processes affects mortality

(a) ToE: Mortality and hospitalisation processes



(b) ToE: a multistate/illness-death model



with a baseline hazard, age dependence, $\lambda_{h0}(t)$, unobserved time-invariant characteristics v_h , and observed time-invariant characteristics x and education level e . We assume a Gompertz baseline hazard, which assumes that the hazard increases exponentially with age: $\lambda_{h0}(t) = e^{\alpha_h t + \alpha_{h0}}$. β_x^h captures the impact of exogenous individual characteristics, x on the hospitalisation hazard and β_e^h captures the impact of (possibly endogenous) individual education, e on the hospitalisation hazard.

As the individual is either in hospital or not, the hospitalisation is alternating, and has three possible transitions: admittance, discharge and, (the absorbing state) death. The conditional hazards for the discharge and re-admittance spells also follow MPH models:

$$\theta_d(d^h|x, v_d) = v_d \lambda_{d0}(d) \exp\left(x\beta_x^d + e\beta_e^d\right) \quad (2)$$

$$\theta_r(d^r|x, v_r) = v_r \lambda_{r0}(d) \exp\left(x\beta_x^r + e\beta_e^r\right), \quad (3)$$

with transition specific Weibull baseline hazards $\lambda_{k0}(d) = \alpha_k d^{\alpha_k - 1}$, unobserved time-invariant characteristics v_k , and observed individual characteristics x , where $k \in \{d, r\}$ denotes the hospitalisation state.

The mortality hazard, our main outcome, is also of the MPH form. We allow the mortality hazard to depend on the timing of the hospitalisation process through a direct effect of hospitalisation, captured by $I^h(t)$ and $I^o(t)$, or, indirectly, through correlated unobservable heterogeneity terms:

$$\theta_m(t|t_h, t_d, t_r, x, e, v_m) = v_m \lambda_{m0}(t) \exp\left(x\beta_x^m + e(\beta_e^m + \gamma_{he}I^h(t) + \phi_{oe}I^o(t))\right). \quad (4)$$

The age dependence of the mortality hazard is assumed Gompertz. The Gompertz hazard is known to provide accurate mortality hazards (Gavrilov and Gavrilova, 1991). Our parameters of interests are β_e^m , the effect of the individual education level and the effect of staying in a hospital, γ_{he} , and the effect of hospital experience, ϕ_{oe} (both also depending on the education level of the individual), on the mortality hazard. For identification we set $\beta_1^m = 0$, primary education, as the reference category.

For the sake of parsimoniousness, we assume that each of the unobserved heterogeneity terms is time invariant, remains the same for recurrent durations of the same type, is

independent of observed characteristics x and, we adopt a discrete distribution (which is standard in the "Timing-of-events"-literature), i.e. v has discrete support (V_1, \dots, V_W) , with $V_w = (v_{h,w}, \dots, v_{m,w})$ and $p_w = \Pr(V = V_w)$.³ The likelihood function is given in Appendix A.

The identification of the effect of hospitalisation on the mortality in the "timing-of-events" model hinges on two key assumptions. The first assumption is the **mixed proportional** structure of the hazard rates, as reflected in the hazard structure defined in (1) to (4). This assumption is necessary to distinguish between true duration dependence modelled via the λ_{k0} -functions and dynamic selection because the group of men with early entry in hospital has a different composition than the group with late entry. Similarly, the group of men with long hospital stays is different from the group with much shorter hospital stays. Dynamic selection is taken into account via observed and unobserved heterogeneity. It is well known that in the absence of observed heterogeneity, true duration dependence cannot be distinguished from unobserved heterogeneity (Elbers and Ridder, 1982, Heckman and Singer, 1984).

The second assumption is the no-anticipation assumption, which is defined in terms of the mortality hazard of the potential outcomes, the mortality rate that would be observed if the individual entered hospital at age s (including those who never enter, $s = \infty$), $\theta_m^{(s)}(t|\cdot)$:

No anticipation Assumption: $\theta_m^{(s_1)}(t|\cdot) = \theta_m^{(s_2)}(t|\cdot)$ for all $t \leq \min(s_1, s_2)$

In other words, we assume that individuals do not anticipate entering hospital for CVD by dying before the anticipated event would occur. This assumption does not hold if individuals died because they know they will enter hospital for CVD. Although we think this is rather unlikely (but untestable with our data) we are cautious in using a causal interpretation of the obtained CVD hospitalisation effects. Still, even if the no-anticipation assumption does not hold the timing-of-events method corrects for possible endogeneity of the hospitalisation processes.

The identification of the model framework is proven and discussed at length in Abbring and van den Berg (2003). To provide some intuition, first note that the data can be broken into three parts: (*i*) a competing risk part for the duration until either a recruit is admitted

³To assure that the probability is between zero and one we estimate q_w with $p_w = e^{q_w} / (1 + \sum e^{q_j})$.

to hospital for CVD or dies, whichever comes first, (ii) a competing risk part for the duration until either discharge from hospital or until death, and (iii) the residual duration from the moment of hospital experience until death. From Heckman and Honoré (1989), it follows that under general conditions the whole model except for γ_{he} and ϕ_{oe} is identified from the data corresponding to the competing risk parts. Subsequently, γ_{he} and ϕ_{oe} are identified from the data corresponding to part (iii) of the model.

To clarify further what drives the identification of γ_{he} and ϕ_{oe} , consider individuals who enter hospital for CVD at age t . The natural control group consists of individuals who have not been in hospital at t . A necessary condition for a meaningful comparison of these groups is that there is some randomization in the hospital admittance at t . The duration model framework allows for such a randomization because it specifies assignment by the way of the rate of entering hospital. In addition, we have to deal with the selection issue that the unobserved heterogeneity distribution is different between the hospitalised and unhospitalised groups at t . This is handled by exploiting the information in the data on what happened to individuals who entered hospital or died before age t .

Another way to look at this is to note that the timing of the consecutive events of admittance to and discharge from hospital is informative on the presence of the causal effect of a hospitalisation. If CVD hospitalisations are often followed very quickly by death, then this indicates a positive hospitalisation effect. The selection effect does not give rise to the same type of quick succession of events.

The no anticipation assumption has been used throughout the empirical literature on dynamic treatment effects, e.g. van den Berg et al. (2004), Crépon et al. (2018), including health economics (see a.o. van den Berg et al. (2011), van Ours and Williams (2012), Baert et al. (2018)). Moreover, it is an implicit assumption in some standard static approaches, such as the Difference-in-Difference (DiD) and the Synthetic Control approaches. Note that identification of the model does not require exclusion restrictions or assumptions on the functional form of the baseline hazards or the distribution of the unobservables $\{v_m, v_r, v_d, v_h\}$. Identification of the parameters of interest γ_{he} and ϕ_{oe} comes from the extent to which CVD hospitalisation is followed closely by death, regardless of the other mortality determinants. Monte Carlo evidence has shown that these models are reliably estimated with samples where as few as a couple of hundred individuals are “treated” (in our case: experience a CVD hospitalisation);

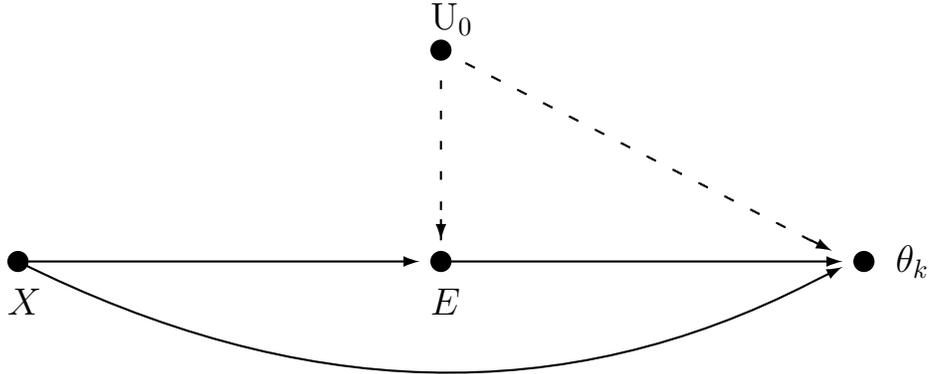


Figure 4: Directed acyclic graph of selection through X and U_0 and the impact of education E on the hazard rates θ_k (for $k = \{h, r, d, m\}$).

see e.g. Gaure et al. (2007).

3.2 Accounting for selective educational attainment

However, mortality and the hospitalisation process may be influenced by factors that also determine the education choice. This may render education a selective choice and makes it endogenous to mortality and hospitalisation. The timing-of-events method still fails to correct for possible endogeneity of educational attainment.

We follow a propensity score method to account for selection on observed characteristics and estimate the effect of education on the hospitalisation rates and on the mortality rate. Figure 4 provides a graphical illustration of the relationship between observed characteristics, education and the hazards from the timing-of-events model θ using a directed acyclic graph, where each arrow represents a causal path (Pearl, 2000, 2012). It states that observed early childhood characteristics X , such as parental background and intelligence, influence the education choice E , and the hazards of the hospitalisation and mortality processes. Possible unmeasured childhood (pre-age 18) factors, U_0 , may also influence both the education choice and the hazards.

We base our (generalised) propensity scores on a multistage sequential model of educational choice developed by Cameron and Heckman (2001) with five education levels $e =$

1, . . . , 5 (see also (Heckman et al., 2018b,a)). People start their educational career in primary school ($e = 1$) and choose if they wish to finish some secondary school ($e = 2$). Importantly, the set of future choices available to them depends on their earlier educational choices. If people choose to finish some secondary school, they have the choice to graduate secondary school ($e = 3$), if they graduate they have the choice to enroll in post-secondary schooling ($e = 4$), if they enroll, they have the choice to graduate from post-secondary schooling, if they graduate post-secondary schooling they have the choice to enroll in higher education ($e = 5$) and, if they enroll they have the choice to graduate from higher education.

We assume a Probit model for the probability of attaining a higher education level conditional on having obtained the (previous) lower level, based on the sequential model for educational attainment. The propensity score of choosing education level e using the sequential probit assumption is:

$$\Pr(E = e) = \Phi(-\xi_{e+1}X)^{I(e < 5)} \prod_{1 < j \leq e} \Phi(\xi_j X) \quad (5)$$

with $\Phi(\cdot)$ is the standard Normal cumulative distribution. In the propensity scores we control for maternal socio-economic status, paternal education, maternal and paternal age at birth, birth order, IQ and psychological assessment (all obtained at the military examination, when the individuals are 18 years of age), see Table C.6 in Appendix C. We check whether the propensity score is able to balance the distribution of all included variables in all education groups by calculating the standardized bias, or normalised difference in means, see Table B.3 in Appendix B. The overlap, or **common support assumption** requires that the propensity score is bounded away from zero and one, see Figure D.1-D.1 in Appendix D.

First, we discuss the assumptions, common in the potential outcomes literature that uses propensity score methods, to identify the impact of education on the mortality outcomes. We use potential hazard rates, the hazard rate that would be observed if the individual had obtained education level $e = 1, \dots, 5$, $\theta_k(t|e)$ for $k = \{h, r, d, m\}$. We observe pre-treatment (educational level) covariates X that influence the education choice.

Unconfoundedness Assumption: $\theta_k(t|e) \perp E|X$ for $e = 1, \dots, 5$ and $k \in \{h, r, d, m\}$

where \perp denotes independence. The unconfoundedness assumption (Rubin, 1974, Rosenbaum and Rubin, 1983) asserts that, conditional on covariates X , the education level is independent of the potential outcomes. This assumption requires that all variables that affect both the hazards and the education choice are observed. Thus, there is no U_0 in Figure 4. Note that this does not imply that we assume all relevant covariates are observed. Any missing factor is allowed to influence either the hazards or the education choice, not both. We check the robustness of our estimates to this, rather strong, unconfoundedness assumption by assessing to what extent the estimates are robust to violations of this assumption in Section 4.3.

Rosenbaum and Rubin (1983) show that if the potential outcomes are independent of treatment (in our case education) conditional on covariates X , they are also independent of treatment conditional on the propensity score. Hence if unconfoundedness holds, all biases due to observable covariates can be removed by conditioning on the propensity score (Imbens, 2004). The average effects can be estimated by matching or weighting on the propensity score. We follow an inverse propensity weighting method to account for this endogeneity (Hirano et al., 2003). Inverse probability weighting based on the propensity score, or inverse propensity weighting (IPW), creates a synthetic sample in which the educational attainment is independent of the included covariates. The synthetic sample is the result of assigning to each individual a weight that is proportional to the inverse of their propensity score. Note that the only relation between the assumptions for the propensity score and the assumptions for the timing-of-events model is that the covariates used to calculate the propensity score are independent of the unobserved heterogeneity of the timing-of-events model. This implies that we assume that the unobserved heterogeneity factors in the hazards do not influence the education choice.

Based on the generalised propensity scores we re-estimate the Timing-of-events models using a re-weighted pseudo-population based on inverse generalised propensity score weighting (IPW-ToE) (Frölich, 2004, Feng et al., 2012), see Bijwaard and Jones (2019) for an application of (a single valued) IPW method for a Mixed Proportional Hazard mortality model.

Misspecification of the propensity score will generally produce bias. Rotnitzky and Robins (1995) point out that if either the regression adjustment or the propensity score is correctly specified the resulting estimator will be consistent. To account for this, we also use a doubly robust estimator, that includes the covariates both in the propensity score and in a regression

adjustment.

Although the unconfoundedness assumption is not directly testable and clearly a strong assumption, it may be a reasonable approximation. Bijwaard and Jones (2019) have shown that intelligence, as measured by an IQ-test, is a principal source of education selection and including this information in the propensity score is robust to possible unconfoundedness violation. The literature has developed a few ways to address violation of the unconfoundedness assumption, e.g. Imbens (2003), Nannicini (2007), Ichino et al. (2008), Bijwaard and Jones (2019). In Section 4.3 we discuss some sensitivity analyses in detail.

4 Results

We estimate both a timing-of-events model without accounting for endogeneity of education (ToE-model) and a timing-of-events model that accounts for this endogeneity through inverse propensity weighting (IPW-ToE-model), as described in the previous Section. Table 2 presents the estimated effect on the mortality hazard of education and CVD hospitalisation by education for both models.⁴ We observe a clear educational gradient. The mortality rate for men with some secondary education is 44% ($= 1 - e^{-0.585}$) lower than the mortality rate for men with only primary education. For men with full secondary education the mortality rate is 60% lower, for men with post-secondary education 70% lower and, for men with higher education 73% lower.

For all education levels we find that the mortality rate is higher in hospital and also for those who have experienced CVD hospitalisation. For both hospitalisation effects we find a clear educational gradient. For example, consider the difference in the estimated effects of CVD hospitalisation while in hospital on the mortality rate for men with primary education versus men with higher education. While for men with primary education the mortality rate in hospital is 68 ($= e^{4.221}$) times higher, for men with higher education it is 39 times higher (compared to an individual with higher education not in hospital and without hospital experience, or 11 ($= e^{3.66} \times e^{-1.303}$) times higher compared to an individual with primary education not in hospital and without hospital experience). Similarly, the mortality rate for men with higher education is ‘only’ 2 times higher when they have experienced CVD

⁴A full overview of estimated coefficients can be found in Table C.1 to Table C.4 in Appendix C.

hospitalisation (or even 44% lower compared to the mortality rate of men with primary education without hospitalisation experience).

Accounting for education endogeneity through IPW, reported in the second panel of the table, slightly affects the estimated coefficients of education on the mortality rate. The estimated educational impact on mortality for the two highest education groups is smaller when using the IPW correction. Thus, this reduces the educational gradient in mortality. Accounting for educational endogeneity also changes the estimated impact of hospitalisation. On the one hand, for the higher education groups the IPW-correction (slightly) increases the estimated impact of being in hospital on the mortality rate and, therefore, decreasing the educational gradient in the hospitalisation effect. On the other hand, the IPW correction reduces, for all education groups, the impact of hospital experience on the mortality rate.

Table 2: Impact of hospitalisation and education on mortality hazard

		Education level ^a				
		(1)	(2)	(3)	(4)	(5)
ToE-model	education	—	−0.587** (0.016)	−0.928** (0.025)	−1.300** (0.028)	−1.443** (0.027)
	in hospital	4.213** (0.055)	4.016** (0.053)	3.854** (0.085)	3.642** (0.096)	3.546** (0.106)
	hospital experience	1.830** (0.108)	1.620** (0.108)	1.400** (0.114)	1.096** (0.116)	0.928** (0.117)
IPW-ToE-model	education	—	−0.585** (0.016)	−0.863** (0.025)	−1.216** (0.027)	−1.303** (0.027)
	in hospital	4.221** (0.060)	4.038** (0.055)	3.966** (0.084)	3.718** (0.090)	3.660** (0.094)
	hospital experience	1.546** (0.110)	1.370** (0.109)	1.236** (0.115)	0.837** (0.117)	0.730** (0.118)

^a (1) primary education; (2) Some Secondary education; (3) Full secondary education; (4) Post-secondary education; (5) University or PhD. + $p < 0.05$, ** $p < 0.01$.

4.1 Implied Educational gain

In a single risk (e.g. only mortality) proportional hazard model, the educational coefficients have a clear interpretation as providing the proportionality factor of the hazard for a given education level (or conditionally on the unobserved heterogeneity for a mixed proportional hazard rate model). However, in our timing-of-events model, when the effect of the education is also modelled in the hospitalisation process, for the hazard of first hospitalisation (1), the discharge hazard of discharge (2) and, the admission hazard (3), this does not hold. The reason is that (total) survival not only depends on the mortality hazard, but also on the hospitalisation hazards. Hence, the impact of a specific education level on survival not only depends on the parameter value of this education level on the mortality rate, but also on the effect of this education level on all hospitalisation hazards. Admission to and discharge from hospital also show a clear educational gradient, see Table C.1 to Table C.3 in Appendix C. Including interactions in the hazards between the education level and the status of the hospitalisation process further complicates interpretation. As a result the reported coefficients of education and hospitalisation in Table 2 are rather difficult to interpret. The measures we derive in this section provide information with clear interpretation on the impact of improving education (by one level) on survival and the length of life, while accounting for the age dependence of mortality, the impact of education on the mortality and, the educational differences in the hospitalisation process.

We use counterfactual simulations to assess the educational gain for two such measurements: (1) the educational gain in the survival probability up till age 63 and (2) the educational gain in the number of months lost due to early mortality before that age. The first educational gain estimates the increase in the probability of survival from age 18 (age at military examination) to age 63, the highest age obtained at the end of our observation window, if the men had had, contrary to what is observed, (one level) higher education. For example, we compare the survival probability (implied by the estimated IPW-ToE model) of men with primary education and what their survival probability would have been if they had had some secondary education instead. The estimated increase in the survival probability is the educational gain. The second educational gain measure is based on the months-lost concept and is defined as the difference (decrease) in the expected number of months lost for

men if they had improved their education with one level. For example, a men who dies at age 60 (for CVD or other causes), has lost 36 months of his possible life until age 63. For example, for men with full secondary education this educational gain shows how much less months these men would have lost (up till age 63) if they had had post-secondary education instead.

For both outcomes we decompose the total educational gains into a *direct effect* and an *indirect effect*. The direct effect is the educational gain not running through CVD hospitalisation, thus capturing the effect of education on survival or months-lost through other factors that influence mortality. The indirect effect is the educational gain running through changes in CVD hospitalisation (due to education), e.g. how much would the survival probability till age 63 increase when men with primary education had the hospitalisation process of men with some secondary education.

The simulation is based on a synthetic cohort of individuals, all entering at age 18. The synthetic cohort consists of 10,000 men not in hospital, for which the distribution equals the observed distribution at the military examination (ignoring the educational distribution). For each simulation round (of 100 rounds), we draw a vector of parameter estimates using the fact that the estimated coefficients of the IPW-ToE model are normally distributed around the point estimates with a variance-covariance matrix equal to the estimated one. Then, on a monthly basis, we simulate the transitions for each member of the synthetic cohort using the implied transition intensities derived from the simulated parameter values. Originally, the men can only enter hospital (for CVD) or die and we are using the (first) hospitalisation and mortality hazards to simulate these events. When a simulated man enters hospital, we continue to use the discharge hazard from hospital (and adjust the mortality hazard), and similarly for a man with simulated hospital experience we simulate possible re-admission (and adjust the mortality hazard). In this fashion we obtain, for each simulation round, the simulated hospitalisation and mortality process for each of the 10,000 men in the synthetic cohort. Of course, this also includes (many) men who do not experience (are simulated to) any hospitalisation and/or live beyond age 63.

The simulated survival outcome is the average (averaged over the rounds and the synthetic cohort) percentage of simulated men that survives till age 63. The simulated months lost outcome is the average month lost (months before age 63) for each simulated man. Note that

when a (simulated) man survives beyond age 63, the survival is censored at age 63, and no months alive are lost.

For the counterfactual simulations we impose a given education level for the hospitalisation hazards and the mortality hazards. Let $Y(e_1, e_2)$ be the simulated average outcome (survival probability or months lost) for men with a mortality rate given education level e_1 and hospitalisation hazards given education level e_2 , then the educational gain for improving education from e to $e + 1$ is:

$$G(e, e + 1) = Y(e + 1, e + 1) - Y(e, e) \tag{6}$$

$$= \left[Y(e + 1, e + 1) - Y(e, e + 1) \right] + \left[Y(e, e + 1) - Y(e, e) \right] \tag{7}$$

with the direct effect, the educational gain not running through CVD hospitalisation, in the first term in brackets of (6) and the indirect effect, the educational gain running through changes in CVD hospitalisation, in the second term in brackets.

Table 3 reports the estimated educational gains by education level. Improving education (with one level) would lead to 2% to 5.5%-point increase in the survival probability from age 18 to age 63 and 2 to 9 months longer expected life (till age 63). The educational gains (for both measures) are clearly the largest for men with the lowest education level, men with primary education. Except for an improvement of education for men with post-secondary education the direct effect of educational improvement, i.e. from other factors affected by education than CVD hospitalisation, is statistically significant and the main source of the educational gain. Only for men who would improve their education level from some to full secondary education or from post-secondary to university the indirect effect of CVD hospitalisation on the educational gain in the survival probability is statistically significant. In fact, for the latter men the educational gain in the survival probability is mostly due to differences in the hospitalisation process between the two education levels.

For the educational gain in months-lost, the differences in the educational parameters of the hospitalisation hazards do not significantly contribute to the decrease in months lost when a man would have had a higher education level. The reason why the educational differences in the CVD hospitalisation process do not lead to significant educational gain in months alive is probably due to the early censoring age of 63. We expect that we would have found significant

educational gains if we could have followed the men until they had reached a higher age.

Table 3: Implied educational gain and its decomposition into a direct and an indirect effect

	Education level ^a			
	(2)	(3)	(4)	(5)
	survival probability till age 63			
Total effect	0.055** (0.007)	0.031** (0.007)	0.030** (0.007)	0.019** (0.007)
Direct effect ^b	0.049** (0.007)	0.017+ (0.007)	0.026** (0.007)	0.005 (0.007)
Indirect effect ^b	0.005 (0.007)	0.014** (0.005)	0.004 (0.005)	0.013** (0.005)
	months lost 18–63			
Total effect	9.08** (1.19)	4.33** (0.98)	4.00** (0.97)	2.02+ (0.89)
Direct effect ^b	8.53** (1.14)	2.75+ (1.08)	3.56** (1.04)	0.74 (0.89)
Indirect effect ^b	0.54 (1.02)	1.58 (0.86)	0.43 (0.80)	1.27 (0.78)

^a (2) primary education to Some Secondary education; (3) Some Secondary education to Full secondary education; (4) Full secondary education to Post-secondary education; (5) Post-secondary education to University or PhD. + $p < 0.05$, ** $p < 0.01$.

^b Indirect effect: effect of education running through CVD hospitalisation; Direct effect: effect of education running through other (not CVD hospitalisation) factors

4.2 Robustness

In this section, we present a couple of robustness checks. First, we investigate whether removing the possibility of reverse causation changes the estimated impact of education (and hospitalisation) on mortality. Reverse causation might occur as education influences both psychological fitness and intelligence measured at the military examination. A couple of studies have shown that additional education improves intelligence (Falch and Sandgren Massih, 2011, Banks and Mazzonna, 2012, Schneeweis et al., 2014, Carlsson et al., 2015, Dahmann, 2017). In that case, intelligence is a mediator in the causal path from education to health (Bijwaard and Jones, 2019). Ideally, we would have multiple measurements of the (development) of intelligence over the life cycle, to account for both the selection and mediation of intelligence in the causal path from education to mortality. However, in our data, we only observe intelligence at late adolescence (during the military examination) when measured IQ

can be either the result of the attained education or a proxy of early childhood intelligence which influences education choice. Similar reasoning holds for psychological fitness.

To account for this possible ‘reverse causation’ we also estimate models without psychological assessment or IQ measurement in the propensity score. The results in the first panel of Table 4 reveal that when leaving out these measurements from the propensity score the estimated coefficients of education and hospitalisation for CVD in the mortality hazard change significantly. Both the direct educational gradient (except for some secondary education) and the “hospital experience”-effect and its implied educational gradient become (significantly) larger. However, this would violate the unconfoundedness assumption as these measurements influence both the education choice and the hospitalisation and mortality process. From the estimated coefficients of IQ and psychological assessment, given in Table C.5 and Table C.6 in Appendix C, it is clear that these measurements are very important for both education attainment and the hospitalisation and mortality hazards. We, therefore, prefer the model with IQ and psychological assessment measurements included in the propensity score.

Second, throughout we have assumed that the propensity scores are estimated consistently. Misspecification of the propensity score will generally produce biases. Rotnitzky and Robins (1995) point out that if either the regression adjustment or the propensity score is correctly specified the resulting estimator will be consistent. Thus, to improve the robustness of the proposed methodology we estimate a doubly robust estimator of the model, which also includes a regression adjustment (using the same control variables as included in the propensity score). The results in the second panel of Table 4 indicate that these doubly robust estimates are very similar to the original estimates, reported in the second panel of Table 2.

Third, the association between education and hospitalisation and mortality could also stem from another type of ‘reverse causality’, in which childhood ill-health constrains educational attainment (Behrman and Rosenzweig, 2004, Case et al., 2005). To avoid such reverse causation, we had left out any health measurements from the propensity score. To test whether the estimated effects are affected by this choice we estimate a, double robust, IPW-ToE model that includes health measurements at age 18 available in the data. These additional health measurements include the height (also height-squared), the BMI (= weight in kg/(length in meters)²), BMI-squared, systolic and diastolic blood pressure (and both squared) and grip strength. The results in the third panel of Table 4 indicate

that the estimated education and hospitalisation effects on the mortality rate including these health measurements into both the propensity score and as control variables in the hazards are very similar to the original estimates.

Table 4: Impact of hospitalisation and education on mortality hazard IPW, robustness

	Education level ^a				
	(1)	(2)	(3)	(4)	(5)
	<i>without IQ or psychological assessment^b</i>				
education	–	–0.597 (0.016)	–0.965** (0.025)	–1.297** (0.027)	–1.453** (0.028)
in hospital	4.221 (0.055)	3.969 (0.053)	3.793+ (0.086)	3.668 (0.094)	3.327** (0.108)
hospital experience	1.966** (0.107)	1.713** (0.107)	1.485+ (0.113)	1.159** (0.115)	1.045** (0.116)
	<i>doubly robust^c</i>				
education	–	–0.597 (0.016)	–0.873 (0.025)	–1.238 (0.027)	–1.325 (0.027)
in hospital	4.227 (0.057)	4.002 (0.052)	4.023 (0.082)	3.692 (0.089)	3.614 (0.093)
hospital experience	1.637 (0.108)	1.423 (0.107)	1.321 (0.113)	0.915 (0.115)	0.796 (0.116)
	<i>doubly robust with Health measurements^d</i>				
education	–	–0.604 (0.016)	–0.868 (0.025)	–1.241 (0.027)	–1.398 (0.032)
in hospital	4.174 (0.054)	3.955 (0.048)	3.914 (0.081)	3.658 (0.084)	3.306+ (0.116)
hospital experience	1.572 (0.107)	1.356 (0.107)	1.299 (0.112)	0.922 (0.114)	0.664 (0.120)

^a (1) primary education; (2) Some Secondary education; (3) Full secondary education; (4) Post-secondary education; (5) University or PhD.

^b Without IQ or psychological assessment as control variables in the propensity scores.

^c Including all the covariates both in the propensity scores and all the hazards.

^d Doubly robust with additional health measurements included: height, height-squared, BMI, BMI-squared, systolic and diastolic blood pressure (and squared) and grip strength. Significance of the difference compared to the basic results in the second panel of Table 2

+ $p < 0.05$, ** $p < 0.01$.

In Bijwaard et al. (2017), who also use data on Swedish recruits (more cohorts, but no hospitalisation data), it is argued that the men in the lowest and the highest educational groups differ too much in their observed background characteristics, which causes severe overlap problems in the propensity score. To address this overlap problem, they estimate separate propensity scores of attaining a higher educational level through pairwise comparisons of ad-

adjacent educational levels. They did not consider the sequential probit model we assume for the educational attainment. As a fourth robustness check, we investigate whether using only adjacent education levels would lead to different results. To this end we estimate an IPW-ToE model for each pair of adjacent education levels separately: primary vs some secondary education; some secondary vs full secondary education; full secondary vs post-secondary education and; post-secondary vs university education. The estimated educational effects from these models, in Table B.4 in Appendix B, are not directly comparable to the original results and are, therefore, deferred to the appendix. The reason is that for each pair of education levels the reference group changes. We focus on the implied educational gains instead and compare them with the original educational gains reported in Table 3. These educational gains, reported in Table 5, do not differ significantly from the original educational gains.

Table 5: Implied educational gain and its decomposition into a direct and an indirect effect, adjacent educations model

	Education level ^a			
	(2)	(3)	(4)	(5)
	survival probability till age 63			
Total effect	0.044 (0.007)	0.030 (0.008)	0.027 (0.009)	0.018 (0.008)
Direct effect ^b	0.038 (0.008)	0.016 (0.008)	0.021 (0.008)	0.007 (0.008)
Indirect effect ^b	0.005 (0.007)	0.013 (0.009)	0.006 (0.008)	0.011 (0.008)
	months lost 18–63			
Total effect	7.38 (1.18)	3.85 (0.95)	3.05 (0.93)	1.80 (0.82)
Direct effect ^b	6.94 (1.26)	2.39 (1.01)	2.52 (0.79)	0.79 (0.90)
Indirect effect ^b	0.44 (1.15)	1.46 (1.11)	0.53 (0.83)	1.01 (0.77)

^a (2) primary education to Some Secondary education; (3) Some Secondary education to Full secondary education; (4) Full secondary education to Post-secondary education; (5) Post-secondary education to University or PhD.

^b Indirect effect: effect of education running through CVD hospitalisation; Direct effect: effect of education running through other (not CVD hospitalisation) factors.

Significance of the difference compared to the basic results in Table 3

⁺ $p < 0.05$, ^{**} $p < 0.01$.

4.3 Sensitivity

The critical assumption in propensity score weighting is that of no selection on unobservables. To test the sensitivity of the estimates to this unconfoundedness assumption we build on the sensitivity analyses of Nannicini (2007), Ichino et al. (2008) and, in particular, Bijwaard and Jones (2019) and extend these analyses to the Timing-of-events model with IPW. The two main assumptions for these sensitivity analyses are that the possible unobserved confounding factors can be summarised in a binary variable, U , and that the unconfoundedness assumption holds conditional on X and the additional variable U , i.e. $\theta_k(t|e) \perp D|X, U$ (for $k \in \{h, r, d, m\}$). Given the values of the probabilities that characterize the distribution of U we can simulate a value of the unobserved confounding factor for each individual and re-estimate the IPW-ToE, including the simulated U in the propensity score. A sensible choice for the probabilities of the distribution of U is to mimic the outcome distribution implied by the included covariates (Nannicini, 2007). Although the outcomes are continuous (time till hospitalisation or death) they have a natural binary transformation using the censoring indicator(s) (Bijwaard and Jones, 2019). The censoring indicators are $\delta_M = 1$ when the individual is still alive at the end of the observation period and $\delta_H = 1$ when the individual still has no CVD hospitalisation experience at the end of the observation period. Of course, the censoring distribution also depends on the education level. With five education levels and two times two censor indicator values we have 20 probabilities characterising the distribution of U with

$$p_{emh} = \Pr(U = 1|E = e, \delta_M = m, \delta_H = h, X) = \Pr(U = 1|E = e, \delta_M = m, \delta_H = h) \quad (8)$$

for $m, h = 0, 1$ and $e = 1, \dots, 5$.

A measure of how the different configurations of p_{emh} , chosen to simulate U , translate into associations of U with mortality is $\omega_M(e)$, the coefficient of U in a Mixed Proportional Hazard model for the mortality for each education level ($E = e$) using U and X as covariates. We call this coefficient the ‘mortality effect’. The measures of the association between U and first (CVD) hospitalisation, hospitalisation discharge and re-admission to hospital are $\omega_H(e)$, $\omega_D(e)$ and $\omega_R(e)$ respectively, all obtained from estimating a MPH model on the relevant duration (time till first hospitalisation, time to discharge or time to re-admittance after discharge).

We call these coefficients the ‘first hospitalisation effect’, the ‘discharge effect’ and, the re-admittance effect’. A measure of the effect of U on the relative probability to have chosen education level e is $\xi(e)$, which is the coefficient of U in a sequential probit model of choosing education ($E = e$) using U and X as covariates. In line with Bijwaard and Jones (2019) we call this coefficient the ‘selection effect’.

The probability values of the distribution for U are chosen so that they mimic the distribution for each included binary variable. For example, consider the probability that an individual with full secondary education, $e = 3$, has the highest IQ score, 9. Then, p_{300} is the probability for those individuals who died and experienced CVD hospitalisation before the end of the observation period, p_{310} is the probability for those individuals who survived till the end with hospital experience, p_{301} is the probability for those individuals who died before the end without hospital experience, and p_{311} is the probability for those individuals who survived without hospital experience till the end. For each probability configuration, p_{emh} , of U we repeat the simulation of U , the estimation of the mortality-, hospitalisation- and, selection effects $M = 100$ times and obtain the average of these 100 simulations. The total variance of these averages can be estimated from (see Ichino et al. (2008)):

$$\text{Var}_f = \frac{1}{M} \sum_{m=1}^M s_m^2 + \frac{M-1}{M(M-1)} \sum_{m=1}^M (\hat{f}_m - \bar{f})^2 \quad (9)$$

with $f \in \{\omega, \xi\}$ of each pairwise education comparison, \hat{f}_m is the estimated f in each simulation sample m and s_m^2 is its estimated variance.

For each probability configuration, p_{emh} and each simulated U we re-estimate the IPW-ToE model including U in the propensity score to obtain the education and hospitalisation impact on the mortality hazard, γ_{he} and ϕ_{oe} . Again, for each probability configuration the average impact and its variance, using (9), is calculated.

From the estimated parameters in the doubly robust IPW-ToE model (Table C.5 and Table C.6 in Appendix C) we see that IQ is the most important control variable, influencing education choice, the hospitalisation process and, mortality. We, therefore, focus on the results of the sensitivity analysis when assuming U mimics the observed distribution of the IQ-measurements, i.e. the observed education choice and censoring probabilities are equal to the observed education choice and censoring prevalence for individuals with a given IQ level.

We find indeed that the simulated U 's that mimic the distribution of the IQ-values lead to high and significant mortality, first hospitalisation and selection effects, see Table B.5 in Appendix B.⁵ Table 6 reports the simulated impact of education and hospitalisation on the mortality hazard for including these simulated U 's in the propensity score. We find the largest changes in our IPW-ToE estimates when U mimics the education-hospitalisation-mortality distribution of those with the lowest IQ, global IQ = 1. These differences are, however, not statistically significant.⁶ This seems to indicate that the applied propensity score adequately accounts for the endogeneity of education.

⁵The estimated mortality-, hospitalisation- and, selection effects for the simulated U 's that mimic the distribution of the other control variables are given in Table B.6 and Table B.7 in Appendix B.

⁶The estimated education and hospitalisation impact in the IPW-ToE model using simulated U 's that mimic the education-hospitalisation-mortality distribution of the other control variables are given in Table B.8 in Appendix B. Again, for these U 's we do not find any statistically significant difference with the estimated original education and hospitalisation impact in Table 2.

Table 6: Sensitivity analysis U based on IQ-levels (IPW-ToE model with U included)

	Educational levels ^a				
	(1)	(2)	(3)	(4)	(5)
	Education effect				
original	–	–0.585**	–0.863**	–1.216**	–1.303**
<i>Global IQ</i> ^b					
1		–0.574	–0.820	–1.212	–1.298
2		–0.589	–0.837	–1.223	–1.317
3		–0.589	–0.853	–1.208	–1.254
4		–0.588	–0.863	–1.222	–1.300
6		–0.575	–0.850	–1.199	–1.292
7		–0.589	–0.856	–1.215	–1.300
8		–0.601	–0.865	–1.219	–1.286
9		–0.594	–0.876	–1.224	–1.300
missing		–0.577	–0.851	–1.206	–1.292
	in hospital effect				
original	4.221**	4.038**	3.966**	3.718**	3.660**
<i>Global IQ</i> ^b					
1	4.231	4.054	4.009	3.672	3.620
2	4.246	4.052	3.983	3.674	3.600
3	4.242	4.053	3.977	3.822	3.673
4	4.230	4.049	3.970	3.726	3.679
6	4.221	4.045	3.976	3.715	3.674
7	4.211	4.034	3.969	3.674	3.689
8	4.219	4.044	3.969	3.690	3.682
9	4.204	4.013	3.953	3.687	3.669
missing	4.215	4.038	3.977	3.730	3.669
	Hospital experience effect				
original	1.546**	1.370**	1.236**	0.837**	0.730**
<i>Global IQ</i> ^b					
1	1.533	1.368	1.287	0.775	0.646
2	1.553	1.367	1.260	0.758	0.661
3	1.531	1.349	1.222	0.773	0.746
4	1.527	1.355	1.212	0.790	0.739
6	1.547	1.382	1.253	0.857	0.745
7	1.542	1.374	1.245	0.838	0.748
8	1.525	1.365	1.226	0.823	0.726
9	1.501	1.306	1.201	0.807	0.713
missing	1.546	1.378	1.259	0.856	0.751

^a (1) primary education; (2) Some Secondary education; (3) Full Secondary education 3 years; (4) post-secondary education; (5) Higher

^b Running from low to high.

Based on adding discrete U to propensity score with probabilities of U from observed probabilities for each covariate. Significance of difference with original ⁺ $p < 0.05$ and ^{**} $p < 0.01$

5 Cause of death

In the previous section we have shown that hospitalisation increases and education decreases mortality. Bijwaard et al. (2017, 2019) have shown that, even after accounting for selective education choice, education is negatively associated with most major causes of death. However, these articles ignored the hospitalisation process. Here, we investigate the impact of education and hospitalisation on cause-specific mortality and how the impact of hospitalisation differs by education, by extending the timing-of-events model with competing causes of death.

The data also contain, for those who died, the cause of death. We aggregated the causes of death into five (the first three reflect death due to CVD) categories: (1) Ischemic Heart Disease ; (2) Stroke; (3) other cardiovascular causes; (4) external (suicide, traffic accidents and homicide) causes of death, and (5) Other causes of death. Table 7 reports the percentage of individuals that died from a particular cause before the end of the observation window.

Table 7: Percentage who died by education level and hospitalisation experience

	Primary	Secondary education Some	Full	Post-secondary (< 3 years)	Higher
	<i>All</i>				
Ischemic Heart Disease	1.1%	0.8%	0.6%	0.4%	0.3%
Stroke	0.3%	0.2%	0.1%	0.1%	0.1%
Other CVD	0.6%	0.4%	0.3%	0.2%	0.2%
external causes	3.3%	1.7%	1.2%	0.7%	0.6%
Other causes	4.6%	3.0%	2.4%	1.7%	1.5%
Total # of death	11,427	11,396	2,894	2,258	2,166
	<i>Never in hospital for CVD</i>				
Ischemic Heart Disease	0.9%	0.5%	0.4%	0.3%	0.2%
Stroke	0.1%	0.1%	0.1%	0.0%	0.0%
Other CVD	0.6%	0.4%	0.3%	0.2%	0.2%
external causes	3.7%	1.8%	1.2%	0.7%	0.6%
Other causes	3.9%	2.4%	1.8%	1.3%	1.2%
# of death	8,446	7,881	2,005	1,572	1,575
	<i>After CVD hospitalisation experience</i>				
Ischemic Heart Disease	2.4%	1.9%	1.5%	1.0%	0.8%
Stroke	0.9%	0.8%	0.6%	0.4%	0.5%
Other CVD	1.6%	1.3%	1.0%	0.8%	0.6%
external causes	1.6%	1.1%	0.9%	0.7%	0.4%
Other causes	7.8%	6.2%	5.4%	4.1%	3.8%
# of death	2,981	3,515	889	686	591

To take the timing of the deaths into account, we also calculated the cumulative incidence

functions, the probability of dying from a specific cause of death before some age, with or without hospitalisation. The (non-parametric) Aalen–Johansen cumulative incidence functions Aalen and Johansen (1978) depicted in Figure D.6 in Appendix D for CVD related causes of death and Figure D.7 in Appendix D for external or other natural causes of death, show again a clear educational gradient in the probability to die from each of the five causes of death. Comparing the cumulative incidence curves with and without hospitalisations we notice two things. First, the shape of the cumulative incidence curves for external causes (including suicide, traffic accidents and homicide) without hospitalisation differs substantially from the cumulative incidence curves for other causes of death without hospitalisation and, second, only the probability to die from cardiovascular diseases increases after hospitalisation for CVD. Of course, some caution in interpreting these figures is that the probability somebody has been in hospital for CVD also increases with age and depends on observed and unobserved individual factors and this is not accounted for in the cumulative incidence functions. We account for such dynamic selection in our timing-of-events model.

We use an extension of the timing-of-events model of Section 3 to cause-specific mortality. Instead of one mortality hazard we have five mortality hazards, one for each cause of death. For each of these hazards we assume a MPH form as in (4). To account for possible endogeneity of the hospitalisation process the unobserved heterogeneity of each cause-specific hazard is possibly correlated with the hospitalisation hazards in (1) to (3) and with the other cause-specific hazards. Just as for the analysis of total mortality we account for possible endogeneity of education by using an inverse propensity weighting (in fact the weights are exactly the same based on the same sequential probit estimation of improving education, see Table C.6 in Appendix C).

The results in Table 8 indicate that all causes of death show a clear educational gradient (except for stroke or other natural causes of death between post-secondary and university/PhD), with the largest educational differences for external causes of death. The results also indicate that, not surprisingly, death due to all CVD causes is elevated by CVD hospitalisation. Death due to stroke and other CVD's are affected the most by hospitalisation for CVD. It seems odd that the mortality for external (up to 12 times higher) and other natural causes (up to 61 times higher) is also higher when an individual is in hospital for CVD. A reason for this might be that when discharge from hospital occurs in the morning, which

Table 8: Impact of CVD hospitalisation and education on Cause-specific mortality hazard (IPW)

		Education level ^a				
		(1)	(2)	(3)	(4)	(5)
IHD	education	–	–0.363** (0.049)	–0.535** (0.072)	–0.832** (0.078)	–1.107** (0.082)
	in hospital	4.630** (0.138)	4.421** (0.122)	3.803** (0.274)	3.904** (0.257)	3.334** (0.333)
	hospital experience	4.370** (0.287)	4.292** (0.285)	4.184** (0.296)	3.653** (0.303)	3.954** (0.298)
stroke	education	–	–0.183 (0.096)	–0.724** (0.158)	–1.316** (0.193)	–0.826** (0.150)
	in hospital	6.671** (0.154)	6.405** (0.147)	6.148** (0.236)	5.914** (0.254)	5.722** (0.280)
	hospital experience	3.925** (0.632)	3.351** (0.634)	3.345** (0.671)	2.306** (0.744)	2.551** (0.721)
other CVD	education	–	–0.416** (0.070)	–0.555** (0.102)	–0.947** (0.114)	–0.939** (0.108)
	in hospital	5.537** (0.135)	5.304** (0.123)	5.120** (0.211)	4.690** (0.249)	4.673** (0.253)
	hospital experience	3.618** (0.332)	3.685** (0.329)	3.312** (0.347)	3.013** (0.351)	2.542** (0.367)
other natural	education	–	–0.460** (0.024)	–0.734** (0.036)	–1.065** (0.039)	–1.050** (0.037)
	in hospital	4.110** (0.088)	3.974** (0.080)	4.022** (0.112)	3.806** (0.119)	3.791** (0.120)
	hospital experience	2.537** (0.155)	2.348** (0.154)	2.239** (0.161)	1.809** (0.164)	1.750** (0.165)
external	education	–	–0.810** (0.025)	–1.131** (0.043)	–1.514** (0.048)	–1.789** (0.051)
	in hospital	2.547** (0.261)	2.028** (0.272)	2.685** (0.368)	1.872** (0.525)	1.018 (0.870)
	hospital experience	0.564** (0.066)	0.184** (0.063)	0.112 (0.111)	–0.030 (0.112)	–0.756** (0.161)

^a (1) primary education; (2) Some Secondary education; (3) Full secondary education; (4) Post-secondary education; (5) University or PhD. ⁺ $p < 0.05$, ^{**} $p < 0.01$.

would imply in the data that the individual is still in hospital on that day, could still lead to death for other causes during that day (including external causes, such as a traffic accident or a fatal fall). Unfortunately, we cannot account for this as we do not observe the discharge time at the day of discharge. Another reason for this anomaly might be that we only observe the main cause of death. Co-morbidity could lead to death to other natural causes.

The educational gradient in the hospitalisation effect (decreasing impact of hospitalisation with increasing education) is not clearly present for all causes of death, especially when comparing the highest two education levels. Also surprising is, that for men with the highest education level CVD hospitalisation experience decreases mortality due to external causes. This implies that these men have less traffic accidents and/or suicide after CVD hospitalisation.

5.1 Implied Educational gain

The coefficients in a competing risks Timing-of-events model are even more difficult to interpret than the coefficients of the "standard" timing-of-events model, because the impact of a specific education level on death by a specific cause not only depends on the parameter value of this education level on the cause-specific mortality rate, but also on the effects of this education level on all the other causes-of-death hazards (and still also on all hospitalisation hazards).

We, therefore, derive the implied educational gains, similar to the simulation analysis in Section 4.1 based on the estimated timing-of-events model for the cause-specific mortality with IPW. In a competing risk setting for cause specific mortality we can derive the implied educational gain in months lost for each specific cause of death. If a simulated man dies from a specific cause of death, say ischemic heart diseases, the simulated months lost due to this cause for this man is the number of months this man died before age 63. This implies, for this particular man, he has no months lost for the other causes of death.

Table 9 reports the estimated months lost for each cause of death, decomposed into a direct effect, the educational gain not running through CVD hospitalisation, and the indirect effect, the educational gain running through CVD hospitalisation. We do not find any statistically significant educational gradient in the CVD causes of death. Bijwaard et al. (2017) also found small, but statistically significant, educational gains for death due to cardiovascular diseases. Thus, including the CVD hospitalisation process reduces this even further. However, as also mentioned by Bijwaard et al. (2017), these men are still rather young (max 63) to be really affected by CVD.

The largest educational gains (also in line with Bijwaard et al. (2017)) are found for

external causes of death. Men with primary education would gain 5.7 months alive due to lower external death mortality when their education is improved. Men with some secondary or full secondary would gain about one month alive due to lower external death mortality when their education is improved. Men with primary education are the only men who would gain, about 2.5 months, from educational improvement due to reduced mortality due to other natural causes. For all these significant educational gains we only find a direct effect of education, i.e. from other factors affected by education than CVD hospitalisation.

The educational gains in terms of the cumulative incidence functions, i.e. the probability to die from a particular cause as it evolves over age, are depicted in Appendix D. Figure D.6 and Figure D.7 present the implied cumulative incidence functions by cause of death. The total educational gain (decrease in probability) by cause of death is presented in Figure D.9. Again, they clearly show that men with primary education would gain the most from improving education, especially in reducing death to external causes (which is in line with the analyses on the educational gradient in cause of death by Bijwaard et al. (2017, 2019)). Comparing Figure D.10 and Figure D.11 shows that, again, the main source of educational difference is not caused by difference in CVD hospitalisation, but by other factors related to education.

Table 9: Implied educational gain in months lost 18-63 by Cause of death

	Education level ^a			
	(2)	(3)	(4)	(5)
	Ischemic heart disease			
Total effect	0.40 (0.34)	0.19 (0.31)	0.21 (0.25)	0.19 (0.32)
Direct effect ^b	0.37 (0.34)	0.16 (0.32)	0.19 (0.29)	0.16 (0.34)
Indirect effect ^b	0.03 (0.32)	0.03 (0.28)	0.02 (0.24)	0.03 (0.30)
	stroke			
Total effect	0.07 (0.16)	0.21 (0.14)	0.09 (0.13)	-0.05 (0.17)
Direct effect ^b	0.07 (0.16)	0.16 (0.14)	0.08 (0.14)	-0.08 (0.14)
Indirect effect ^b	0.00 (0.01)	0.05 (0.14)	0.02 (0.12)	0.02 (0.14)
	other CVD			
Total effect	0.27 (0.28)	0.10 (0.24)	0.17 (0.25)	-0.02 (0.29)
Direct effect ^b	0.26 (0.29)	0.07 (0.30)	0.15 (0.24)	-0.02 (0.26)
Indirect effect ^b	0.01 (0.28)	0.02 (0.25)	0.02 (0.24)	0.00 (0.26)
	external			
Total effect	5.68** (0.77)	1.32 ⁺ (0.60)	1.06 ⁺ (0.52)	0.55 (0.59)
Direct effect ^b	5.72** (0.80)	1.25 ⁺ (0.63)	1.11 ⁺ (0.55)	0.61 (0.52)
Indirect effect ^b	-0.04 (0.60)	0.07 (0.52)	-0.06 (0.52)	-0.06 (0.42)
	other natural			
Total effect	2.45** (0.82)	1.06 (0.68)	1.00 (0.60)	-0.06 (0.67)
Direct effect ^b	2.38** (0.74)	1.00 (0.71)	0.94 (0.70)	-0.08 (0.64)
Indirect effect ^b	0.07 (0.68)	0.06 (0.70)	0.06 (0.65)	0.02 (0.64)

^a (2) primary education to Some Secondary education; (3) Some Secondary education to Full secondary education; (4) Full secondary education to Post-secondary education; (5) Post-secondary education to University or PhD. ⁺ $p < 0.05$, ^{**} $p < 0.01$.

^b Indirect effect: effect of education running through CVD hospitalisation; Direct effect: effect of education running through other (not CVD hospitalisation) factors

6 Conclusion and discussion

Higher educated individuals are less frequently admitted to hospital for cardiovascular diseases (CVD) and live longer than their lower educated peers. A relevant question is, therefore, whether the educational gradient in mortality can be explained by the educational difference in CVD hospitalisation. A common approach to obtain the impact of both education and CVD hospitalisation on mortality is to estimate a (mixed) proportional hazard model for the mortality hazard. However, viewing the educational level and the hospitalisation as ordinary (exogenous) variables may lead to biased inference of the effect of these variables on the mortality hazard. Any observed relationship between admittance to (or discharge from) hospital and mortality may be caused by unobserved factors that influence both the hospitalisation and mortality. Educational attainment is also very likely to depend on the same observed factors. Such confounding renders education and hospitalisation endogenous in the mortality analysis. We obtain the impact of education and hospitalisation by education on mortality by accounting for both the selection into the hospitalisation process (both admittance and discharge) and the selection into education.

In particular, we estimate the effects of the hospitalisation process on the mortality rate using the “timing-of-events” - method (Abbring and van den Berg, 2003). We control for correlated effects that arise from correlation between unobservables in the hospitalisation and mortality processes. To account for the endogeneity of the education attainment we apply inverse probability weighting (IPW) methods using the propensity score, based on an estimated sequential probit model for educational attainment. Based on the estimated model we calculate the educational gain of improving education in both the survival probability till age 63 and in the number of months lost till age 63. We also decompose these educational gains into an indirect effect, running through changes in the hospitalisation process, and a direct effect due to other factors.

We base our results on Swedish Military Conscription Data (1951-1960), linked to administrative Swedish registers, linked to administrative Swedish registers including information on CVD hospitalisation and death. We have information on about half a million men who are followed from the date of conscription till the end of 2012, or till death. For those men who die we observe the cause of death. Educational level was classified in five categories:

primary education; some secondary education (2 years); full secondary education (3 years); post-secondary education and higher education.

The empirical analyses show a clear educational gradient in mortality, even after accounting for the endogeneity of education through inverse propensity weighting. We also find that mortality is much higher for those in hospital for CVD and for those who have been in hospital. These hospitalisation effects also exhibit a educational gradient, with a decreasing impact of hospitalisation the higher the education level. From the implied educational gains we conclude that men with only primary education would gain the most if they had a higher education level. However, only a small part of this educational gain is running through a change in hospitalisation for CVD. The only exception are men with post-secondary education for whom the expected gain in months lost if they had higher education is mostly attributable to changes in the CVD hospitalisation for the higher educated.

We present a couple of robustness checks. We investigate whether removing the possibility of reverse causation for psychological assessment and IQ measurement changes the estimated impact of education (and hospitalisation) on mortality. The association between education and hospitalisation and mortality could also stem from childhood ill-health which constrains educational attainment. To test whether the estimated effects are affected by the choice to exclude the health measurement from the propensity score (to avoid such reverse causation) we also estimate a model that includes health measurements at age 18 available in the data. We further address the robustness of our results by comparing the estimation results with the results from a doubly robust model, in which observed characteristics are both included in the propensity score and as control variables in the hazards. For comparison with Bijwaard et al. (2017), we investigate whether using only adjacent education levels leads to significantly different educational gains. Only for the model excluding IQ and psychological assessments from the propensity scores we find, slightly, different estimation results. However, leaving out these variables from the propensity score would violate the unconfoundedness assumption as these measurements influence both the education choice and the hospitalisation and mortality process. We, therefore, prefer the model with IQ and psychological assessment measurements included in the propensity score.

The empirical results for the cause-specific mortality analysis reveal that all five causes of death show a clear educational gradient, with the largest educational differences for external

causes of death. They also reveal that, not surprisingly, death due to CVD (IHD, stroke and other CVD) is affected the most by hospitalisation for CVD. The educational gradient in the hospitalisation effect (decreasing impact of hospitalisation with increasing education) is not present for all causes of death, especially when comparing the highest two education levels. The implied educational gains in months lost by cause of death are only significant for external causes and other natural causes. They do not indicate that hospitalisation (indirect effect) plays an important role in explaining the educational gain. We do not find any statistically significant educational gradient in the CVD causes of death.

During the observation period, Sweden had an advanced public health care system providing services independently of individual income. Educational gains through increasing access to health services through higher income seem to be less important in the context of the Swedish health care system, with its broad coverage and access, than in a health-care system such as that in the United States, in which many individuals are not covered by health insurance. However, the role of education in understanding of health information and in changing health behaviour, healthy life-style, is still potentially present. This study provides better insight in the role of CVD hospitalisation, strongly related to healthy life-style, in explaining the educational difference in mortality. Of course, this is only one of the many possible channels that explain this difference.

A limitation of our data, based on military entrance examination, is that we only observe men and no information on women is available. Another limitation is that, although military conscription was mandatory in Sweden, men with severe mental disabilities or severe chronic diseases were exempted from the military examination. Thus, our results only apply to those who had no severe chronic diseases at age 18 and are, therefore, likely an underestimate of the impact of CVD hospitalisation on mortality. Another limitation is that we only observe mortality before the age of 63. In the future, when these men have been followed for a longer period, the educational differences in mortality and the relevance of CVD hospitalisation in explaining this may change as mortality due to CVD plays a larger role.

The main issue is whether we can give our results a causal interpretation. In the literature three different approaches have been employed to examine the causal effects of education on mortality. The first approach exploits changes in compulsory schooling policies as instrumental variables for educational attainment to control for endogeneity. However, a major limitation

of using changes in compulsory schooling to detect educational effects on mortality, is that often only a relatively small part of the population is affected by the laws (Mazumder, 2008, Fletcher, 2015). Another issue with the instrumental variable methods applied in these studies is that they implicitly assume that the compulsory schooling reforms only affect long-term health through their effect on education, ignoring any other contemporary policy changes that may accompany these reforms. Another identification strategy is to use variation in education among siblings, often identical (monozygotic) twins, to distinguish the unobserved factors shared by these siblings. These studies (Behrman et al., 2011, Lundborg, 2013, Næss et al., 2012, Amin et al., 2015) obtain estimates of the impacts of the differences in schooling within a pair of identical twins on their health differences at various schooling levels. Although by using twins it is possible to control for both shared environmental and shared genetic factors, a major shortcoming of twin studies is that they only analyse twins, yet twins are usually not representative of the whole population. Using twins will substantially reduce the statistical power, because only twins with different education levels are analysed. Not only is it rare that twins would have the same cognitive ability, they also experience a large number of non-shared events throughout life, events that may be unobserved and influence both education and mortality (e.g. accidents). A third approach to account for confounding factors is to include them directly into the model (Bijwaard et al., 2015a,b, 2019) A disadvantage of these models is that they impose a rather stringent structure on the relation between education, mortality and the influence of confounding factors. Another limitation is that estimation of these structural models can be very computer intensive if a large data set is available.

The IPW method we (and Bijwaard et al. (2017)) employ also accounts for possible confounding factors, however, without making any structural assumptions on the relation between the confounding factors and hospitalization or mortality. However, the unconfoundedness assumption of no unmeasured confounding influencing both the education choice and the hospitalisation-mortality process is a strong assumption and nonrefutable. Two findings may indicate that our results are close to causal. First, recent research (Bijwaard and Jones, 2019) has shown that intelligence can be considered a principal source of education selection and, that accounting for intelligence (as we do) leaves little room for selection on unobservables when estimating the impact of education on mortality. Second, the sensitivity analyses indicate that including additional (unmeasured) binary variables with large impact on both

the education choice, hospitalisation and mortality do not change our results. However, we are still cautious in claiming causal interpretation of our results.

References

- Aalen, O. and Johansen, S. (1978). An empirical transition matrix for nonhomogeneous markov chains based on censored observations. *Scandinavian Journal of Statistics*, 5:141–150.
- Abbring, J. H. and van den Berg, G. J. (2003). The non-parametric identification of treatment effects in duration models. *Econometrica*, 71:1491–1517.
- Amin, V., Behrman, J. R., and Kohler, H. P. (2015). Schooling has smaller or insignificant effects on adult health in the US than suggested by cross-sectional associations: New estimates using relatively large samples of identical twins. *Social Science & Medicine*, 127:181–189.
- Arendt, J. N. (2008). In sickness and in health—Till education do us part: Education effects on hospitalization. *Economics of Education Review*, 27:161–172.
- Baert, S., van der Klaauw, B., and van Lomwel, G. (2018). The effectiveness of medical and vocational interventions for reducing sick leave of self-employed workers. *Health Economics*, 27(2):e139–e152.
- Banks, J. and Mazzonna, F. (2012). The effect of education on old age cognitive abilities: Evidence from a regression discontinuity design. *The Economic Journal*, 122(560):418–448.
- Behrman, J., Kohler, H., Jensen, V., Pedersen, D., Petersen, I., Bingley, P., and Christensen, K. (2011). Does more schooling reduce hospitalization and delay mortality? New evidence based on Danish twins. *Demography*, 48(4):1347–1375.
- Behrman, J. R. and Rosenzweig, M. R. (2004). Returns to birthweight. *Review of Economics and statistics*, 86(2):586–601.

- Bijwaard, G. E. and Jones, A. M. (2019). An IPW estimator for mediation effects in hazard models: with an application to schooling, cognitive ability and mortality. *Empirical Economics*, 57(1):129–175.
- Bijwaard, G. E., Myrskylä, M., and Tynelius, P. (2019). Education, cognitive ability and cause-specific mortality: A structural approach. *Population Studies*, 73(2):217–232.
- Bijwaard, G. E., Myrskylä, M., Tynelius, P., and Rasmussen, F. (2017). Educational gains in cause-specific mortality: Accounting for cognitive ability and family-level confounders using propensity score weighting. *Social Science & Medicine*, 184:49–56.
- Bijwaard, G. E. and van Kippersluis, H. (2016). Efficiency of health investment: Education or intelligence? *Health Economics*, 25:1056–1072.
- Bijwaard, G. E., van Kippersluis, H., and Veenman, J. (2015a). Education and health: the role of cognitive ability. *Journal of Health Economics*, 42:29–43.
- Bijwaard, G. E., van Poppel, F., Ekamper, P., and Lumey, L. H. (2015b). Gains in life expectancy associated with higher education in men. *PLoS ONE*, 10:e0141200.
- Caliendo, M. and Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*, 22:31–72.
- Cameron, S. V. and Heckman, J. J. (2001). The dynamics of educational attainment for black, hispanic, and white males. *Journal of Political Economy*, 109(3):455–499.
- Carlsson, M., Dahl, G. B., Öckert, B., and Rooth, D.-O. (2015). The effect of schooling on cognitive skills. *The Review of Economics and Statistics*, 97:533–547.
- Case, A., Fertig, A., and Paxson, C. (2005). The lasting impact of childhood health and circumstance. *Journal of Health Economics*, 24:365–389.
- Clark, D. and Royer, H. (2013). The effect of education on adult mortality and health: Evidence from Britain. *American Economic Review*, 103(6):2087–2120.
- Crépon, B., Ferracci, M., Jolivet, G., and van den Berg, G. J. (2018). Information shocks and the empirical evaluation of training programs during unemployment spells. *Journal of Applied Econometrics*, 33(4):594–616.

- Dahmann, S. C. (2017). How does education improve cognitive skills? Instructional time versus timing of instruction. *Labour Economics*, 47:35–47.
- Elbers, C. and Ridder, G. (1982). True and spurious duration dependence: The identifiability of the proportional hazard model. *Review of Economic Studies*, 49:403–410.
- Falch, T. and Sandgren Massih, S. (2011). The effect of education on cognitive ability. *Economic Inquiry*, 49:838–856.
- Feng, P., Zhou, X.-H., Zou, Q.-M., Fan, M.-Y., and Li, Z.-S. (2012). Generalized propensity score for estimating the average treatment effect of multiple treatments. *Statistics in Medicine*, 31:681–697.
- Fletcher, J. M. (2015). New evidence of the effects of education on health in the US: Compulsory schooling laws revisited. *Social Science & Medicine*, 127:101–107.
- Frölich, M. (2004). Programme evaluations with multiple treatments. *Journal of Economic Surveys*, 18:181–224.
- Galama, T., Lleras-Muney, A., and van Kippersluis, H. (2018). The effect of education on health and mortality: A review of experimental and quasi-experimental evidence. In Hamilton, J. and Jones, A., editors, *Oxford Research Encyclopedia of Economics and Finance (Health, Education and Welfare)*. USA: Oxford UP.
- Galobardes, B., Lynch, J. W., and Smith, G. D. (2004). Childhood socioeconomic circumstances and cause-specific mortality in adulthood: Systematic review and interpretation. *Epidemiologic Reviews*, 26(1):7–21.
- Gaure, S., Røed, K., and Zhang, T. (2007). Time and causality: A Monte Carlo assessment of the timing-of-events approach. *Journal of Econometrics*, 141(2):1159–1195.
- Gavrilov, L. A. and Gavrilova, N. S. (1991). *The Biology of Life Span: A Quantitative Approach*. Harwood Academic Publisher, New York.
- Heckman, J. and Singer, B. (1984). The identifiability of the proportional hazard model. *The Review of Economic Studies*, 51(2):231–241.

- Heckman, J. J. and Honoré, B. E. (1989). The identifiability of the competing risks model. *Biometrika*, 76(2):325–330.
- Heckman, J. J., Humphries, J. E., and Veramendi, G. (2018a). The nonmarket benefits of education and ability. *Journal of Human Capital*, 12(2):282–304.
- Heckman, J. J., Humphries, J. E., and Veramendi, G. (2018b). Returns to education: The causal effects of education on earnings, health and smoking. *Journal of Political Economy*, 126(S1):S197–S246.
- Hirano, K., Imbens, G. W., and Ridder, G. (2003). Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica*, 71:1161–1189.
- Hougaard, P. (2000). *Analysis of Multivariate Survival Data*. Springer–Verlag.
- Ichino, A., Mealli, F., and Nannicini, T. (2008). From temporary help jobs to permanent employment: What can we learn from matching estimators and their sensitivity? *Journal of Applied Econometrics*, 23:305–327.
- Imbens, G. W. (2003). Sensitivity to exogeneity assumptions in program evaluation. *American Economic Review*, 93:126–132.
- Imbens, G. W. (2004). Nonparametric estimation of average treatment effects under exogeneity. *The Review of Economics and Statistics*, 86:4–29.
- Kulhánová, I., Hoffmann, R., Eikemo, T. A., Menvielle, G., and J. P. Mackenbach, J. P. (2014). Educational inequalities in mortality by cause of death: First national data for the Netherlands. *International journal of public health*, 59(5):687–696.
- Lager, A. C. J. and Torssander, J. (2012). Causal effect of education on mortality in a quasi-experiment on 1.2 million Swedes. *Proceedings of the National Academy of Sciences*, 109(22):8461–8466.
- Lancaster, T. (1990). *The Econometric Analysis of Transition Data*. Cambridge University Press.
- Lundborg, P. (2013). The health returns to schooling: What can we learn from twins? *Journal of Population Economics*, 26(2):673–701.

- Lundborg, P., Lyttkens, C. H., and Nystedt, P. (2016). The effect of schooling on mortality: New evidence from 50,000 Swedish twins. *Demography*, 53(4):1135—1168.
- Mackenbach, J. P., Stirbu, I., Roskam, A.-J. R., Schaap, M. M., Menvielle, G., Leinsalu, M., and Kunst, A. E. (2008). Socioeconomic inequalities in health in 22 European countries. *New England Journal of Medicine*, 358(23):2468–2481.
- Mazumder, B. (2008). Does education improve health: A reexamination of the evidence from compulsory schooling laws. *Federal Reserve Bank of Chicago Economic Perspectives*, 33(2):2–16.
- McCartney, G., Collins, C., and Mackenzie, M. (2013). What (or who) causes health inequalities: Theories, evidence and implications? *Health Policy*, 113:221—227.
- Meghir, C., Palme, M., and Simeonova, E. (2018). Education, cognition and health: Evidence from a social experiment. *American Economic Journal: Applied Economics*, 10(2):234–256.
- Næss, Ø., Hoff, D. A., Lawlor, D., and Mortensen, L. H. (2012). Education and adult cause-specific mortality—examining the impact of family factors shared by 871 367 Norwegian siblings. *International Journal of Epidemiology*, 41:1683–1691.
- Nannicini, T. (2007). A simulation-based sensitivity analysis for matching estimators. *STATA Journal*, 7:334–350.
- Pearl, J. (2000). *Causality: Models, Reasoning and Inference*. Cambridge University Press, New York.
- Pearl, J. (2012). The mediation formula: A guide to the assessment of causal pathways in nonlinear models. In Berzuini, C., Dawid, P., and Bernardinelli, L., editors, *Causality: Statistical Perspectives and Applications*, pages 151–179. John Wiley, Chichester.
- Putter, H., Fiocco, M., and Geskus, R. B. (2007). Tutorial in biostatistics: Competing risks and multi-state models. *Statistics in Medicine*, 26:2389–2430.
- Rosenbaum, P. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70:41–55.

- Rotnitzky, A. and Robins, J. M. (1995). Semiparametric regression estimation in the presence of dependent censoring. *Biometrika*, 82:805–820.
- Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and non-randomized studies. *Journal of Educational Psychology*, 66:688–701.
- Schneeweis, N., Skirbekk, V., and Winter-Ebmer, R. (2014). Does education improve cognitive performance four decades after school completion? *Demography*, 51(2):619–643.
- Spasojević, J. (2010). Effects of education on adult health in Sweden: Results from a natural experiment. In Slottje, D. and Tchernis, R., editors, *Current Issues in Health Economics (Contributions to Economic Analysis)*, pages 179–199. Emerald Group Publishing Limited, Wst Yorkshire, UK.
- Tansel, A. and Keskin, H. (2017). Education effects on days hospitalized and days out of work by gender: Evidence from Turkey. DP 11210, IZA.
- Van den Berg, G. J. (2001). Duration models: Specification, identification, and multiple duration. In Heckman, J. and Leamer, E., editors, *Handbook of Econometrics, Volume V*, chapter 55, pages 3381–3460. Amsterdam: North-Holland.
- van den Berg, G. J., Lindeboom, M., and Portrait, F. (2011). Conjugal bereavement effects on health and mortality at advanced ages. *Journal of Health Economics*, 30(4):774–794.
- van den Berg, G. J., van der Klaauw, B., and van Ours, J. C. (2004). Punitive sanctions and the transition rate from welfare to work. *Journal of Labor Economics*, 22(1):211–241.
- van Ours, J. C. and Williams, J. (2012). The effects of cannabis use on physical and mental health. *Journal of Health Economics*, 31(4):564–577.
- Xue, X., Cheng, M., and Zhang, W. (2021). Does education really improve health? A meta-analysis. *Journal of Economic Surveys*, 35(1):71–105.

Appendix A Likelihood function

We have data for $i = 1, \dots, n$ male recruits in our observation window. Let K_{id} and K_{ir} denote the number of the discharges and re-admittances out/in a hospital for CVD of individual i . Note that for some individuals $K_{id} = 0$ and $K_{ir} = 0$, i.e an individual who either never entered a hospital or who died in hospital. An important feature of duration data is that for some individuals we only know that he or she survived up to a certain time (often the end of the observation window). In this case an individual is (right) censored and we use the survival function instead of the hazard in the likelihood function. The three indicators Δ_{ik}^d , Δ_{ik}^r and Δ_i^m signal that k^{th} CVD hospitalisation discharge/re-entry or the mortality spell is uncensored. Δ_i^h indicates that the first CVD hospitalisation spell is uncensored. Thus, the likelihood contribution of individual i conditional on the unobserved heterogeneity $v = (v_h, v_r, v_d, v_m)$ is (suppressing dependence on observed characteristics x and education e), in the light of the preceding discussions:

$$\begin{aligned}
 L_i(v) &= \theta_h(t_i^h | \cdot, v_h)^{\Delta_i^h} \exp\left(-\int_0^{t_i^h} \theta_h(\tau | \cdot, v_h) d\tau\right) \\
 &\quad \times \prod_{k=1}^{K_{id}} \left[\theta_d(d_{ik}^h | \cdot, v_d)^{\Delta_{ik}^d} \exp\left(-\int_0^{d_{ik}^h} \theta_d(\tau | \cdot, v_d) d\tau\right) \right]^{I^h(t_{ik}^-)} \\
 &\quad \times \prod_{j=1}^{K_{ir}} \left[\theta_r(d_{ij}^r | \cdot, v_r)^{\Delta_{ij}^r} \exp\left(-\int_0^{d_{ij}^r} \theta_r(\tau | \cdot, v_r) d\tau\right) \right]^{I^o(t_{ij}^-)} \\
 &\quad \times \theta_m(t_i | \cdot, v_m)^{\Delta_i^m} \exp\left(-\int_0^{t_i} \theta_m(\tau | \cdot, v_m) d\tau\right)
 \end{aligned} \tag{A.1}$$

This likelihood naturally separates admittance, discharge, re-admittance and mortality spells, and for each spell allows for censoring. $I^h(t_{ik}^-)$ indicates that the individual is in hospital just before t_{ik} and similarly for $I^o(t_{ij}^-)$. When $K_{id} = 0$ or $K_{ir} = 0$ the relevant term becomes 1. Note that the last, and only the last, hospitalisation spell is censored. This is either because the individual is still alive at the end of the observation period, or has died.

Another feature of duration data is that only individuals are observed having survived up to a certain age. In our case, mortality follow-up is only available from the conscription date, around age 18, onwards. In this case the individuals are left-truncated, and we need to condition on survival up to the age of first observation, $t_0 = 18$. With left-truncated data the

distribution of unobserved heterogeneity among the survivors (up to the left-truncation time) changes. When only individuals are observed that have survived until age t_0 the likelihood contribution is

$$L_i = \int L_i(v) \exp\left(\int_0^{t_0} \theta_m(\tau|\cdot, v_m) d\tau\right) dG(v|T > t_0)$$

with the distribution of the unobserved heterogeneity conditional on survival up to t_0

$$dG(v|T > t_0) = \frac{\exp\left(-\int_0^{t_0} \theta_m(\tau|\cdot, v_m) d\tau\right) dG(v_m, v_h, v_d, v_r)}{\int \exp\left(-\int_0^{t_0} \theta_m(\tau|\cdot, v_m) d\tau\right) dG(v_m, v_h, v_d, v_r)} \quad (\text{A.2})$$

with $G(v_e, v_u, v_m)$ is the joint distribution of the unobserved heterogeneity terms implied by the discussion of v_k .

Appendix B Additional tables

For an IPW method to hold we need to check if the propensity score is able to balance the distribution of all included variables in all education groups. One suitable way to check whether there are still differences is by calculating the standardized bias, or normalised difference in means:

$$100 \cdot \frac{\bar{x}_e - \bar{x}_p}{\sqrt{\text{Var}(x)_p}} \quad (\text{B.1})$$

With $e = 1, \dots, 5$ the education group and p is the whole sample population. Table B.3 shows the percentage bias measure before and after adjusting the data in our sample. They reveal substantial imbalances between those who attained adjacent education levels before accounting for selective education choice. The biases in columns labelled ‘after’ show that these imbalances disappear when we use the inverse propensity weights.

Table B.1: Descriptive statistics men never in hospital for CVD ($N = 437,614$)

	Primary	Secondary education some	Post-secondary full (< 3 years)	Higher	
	<i>SES mother at birth</i>				
non-manual (high)	1%	1%	3%	4%	7%
non-manual (intermediate)	2%	2%	4%	5%	9%
non-manual (low)	13%	19%	29%	33%	41%
Farmers	18%	14%	13%	12%	9%
Skilled workers	50%	48%	39%	35%	24%
Unskilled workers	10%	10%	8%	7%	6%
Not classified	5%	5%	3%	3%	3%
Unknown	1%	1%	1%	1%	1%
	<i>Education father</i>				
Primary (< 9 yrs)	65%	59%	48%	43%	30%
Primary (9–10 yrs)	3%	3%	4%	4%	4%
Secondary education (2 yrs)	11%	15%	17%	17%	16%
Secondary education (3 yrs)	5%	7%	11%	14%	16%
Post-secondary	1%	2%	3%	5%	6%
Higher	1%	2%	5%	6%	18%
Unknown	14%	12%	12%	11%	11%
mother < 20 at birth	10%	9%	7%	5%	4%
father > 40 at birth	15%	13%	12%	12%	11%
birth order	2.3	2.1	2.0	1.9	1.8
global IQ ^a	4.1	4.6	5.7	6.2	6.8
Psychological assessment ^a	4.4	4.8	5.4	5.7	5.9
missing IQ	14%	14%	12%	12%	12%
missing Psychological assessment	15%	15%	12%	12%	12%
# of individuals	93,219	156,123	53,654	61,786	72,832

^a stanine score 1-9 running from low to high.

Table B.2: Descriptive statistics men who experienced CVD hospitalisation ($N = 80,229$)

	Primary	Secondary education some	Post-secondary full (< 3 years)	Higher	
	<i>SES mother at birth</i>				
non-manual (high)	1%	1%	4%	4%	8%
non-manual (intermediate)	2%	3%	5%	6%	10%
non-manual (low)	14%	19%	30%	34%	42%
Farmers	18%	14%	13%	12%	9%
Skilled workers	48%	47%	36%	33%	22%
Unskilled workers	10%	9%	8%	6%	5%
Not classified	5%	5%	3%	3%	3%
Unknown	2%	1%	1%	1%	1%
	<i>Education father</i>				
Primary (< 9 yrs)	62%	57%	44%	40%	27%
Primary (9–10 yrs)	3%	3%	4%	4%	4%
Secondary education (2 yrs)	11%	15%	17%	17%	16%
Secondary education (3 yrs)	5%	8%	12%	15%	15%
Post-secondary	1%	2%	4%	6%	7%
Higher	2%	2%	6%	8%	21%
Unknown	15%	12%	12%	10%	10%
mother < 20 at birth	9%	9%	7%	5%	4%
father > 40 at birth	11%	11%	10%	10%	10%
birth order	2.4	2.2	2.0	1.9	1.8
global IQ ^a	3.9	4.5	5.6	6.1	6.7
Psychological assessment ^a	4.2	4.6	5.3	5.6	5.8
missing IQ	9%	5%	4%	3%	3%
missing Psychological assessment	10%	6%	4%	4%	4%
# of individuals	20,692	30,737	9,358	9,764	9,678

^a stanine score 1-9 running from low to high.

Table B.3: Standardized bias before and after matching, pairwise comparisons

	Educational levels ^a									
	(1)		(2)		(3)		(4)		(5)	
	Before	After	Before	After	Before	After	Before	After	Before	After
<i>Mother's SES</i>										
Unskilled workers	6.41	0.17	3.99	-0.08	-1.99	-0.19	-6.02	0.10	-11.15	0.39
Farmers	12.12	-1.62	1.18	-0.49	-1.89	-0.60	-4.93	-0.11	-13.68	0.73
Non-manual (low)	-25.12	0.91	-13.88	0.48	10.93	-0.82	20.54	-0.26	39.95	-1.00
Non-manual (medium)	-12.21	0.75	-8.49	0.75	2.44	0.09	8.87	-0.02	26.54	-0.06
Non-manual (high)	-10.36	1.97	-9.19	-0.27	4.52	-0.26	4.78	0.06	27.52	-0.30
not classified	3.97	0.30	3.17	-0.10	-3.24	0.25	-4.85	-0.21	-5.99	0.17
missing	1.57	-0.25	-0.47	-0.37	-0.15	-0.16	-1.16	-0.77	0.02	-1.33
<i>Father's education</i>										
less than 9 years	26.27	-1.89	14.54	-0.35	-10.16	0.71	-19.12	0.56	-44.87	2.05
9-10 years	-3.56	-0.18	-1.14	-0.19	3.22	-0.13	3.25	0.23	2.23	0.01
Full secondary	-15.85	0.20	-7.12	-0.48	6.10	-0.23	16.19	-0.64	19.32	-0.74
University < 3 years	-11.29	0.52	-5.75	0.37	3.28	-0.14	10.86	0.08	16.68	-0.36
University ≥ 3 years	-18.55	3.54	-16.12	1.97	1.48	0.02	7.70	0.06	54.31	-0.32
PhD studies	-6.59	1.00	-6.50	2.11	-0.63	0.13	-0.46	0.07	24.69	-0.03
missing	8.60	-0.88	0.06	-0.79	-1.70	-0.90	-5.83	-0.33	-5.66	-2.12

^a (1) primary education to Secondary education (max 2 years); (2) Secondary education (max 2 years) to Secondary education 3 year; (3) Secondary education 3 years to post-secondary education; (4) Post secondary education to University or PhD.

Table B.3: Standardized bias before and after matching (continued)

	Educational levels ^a									
	(1)		(2)		(3)		(4)		(5)	
	Before	After	Before	After	Before	After	Before	After	Before	After
Mother < 20 at birth	7.99	1.40	5.33	0.41	-2.91	0.55	-7.83	0.95	-14.10	3.73
Father > 40 at birth	5.78	-0.77	0.41	-0.67	-2.43	-0.71	-3.78	-0.19	-3.77	-1.84
family size 1	-6.02	0.08	-2.49	0.65	4.85	-0.17	4.62	-0.28	6.24	-0.46
family size 3	-2.94	-0.34	-0.61	-0.37	0.25	-0.21	1.81	-0.34	3.68	-0.17
family size 4	6.94	0.19	2.55	0.04	-3.92	0.41	-5.92	0.43	-7.23	0.16
family size ≥ 5	20.21	0.39	5.73	0.06	-10.28	0.78	-15.00	0.99	-20.02	0.42
birth order 2	-2.49	-0.15	0.18	-0.06	1.13	0.09	1.38	-0.19	0.96	-0.24
birth order 3	5.61	-0.31	1.92	-0.32	-2.94	0.30	-4.64	1.12	-5.83	0.71
birth order 4	8.08	-0.58	3.05	-0.21	-4.20	0.27	-6.88	-0.61	-8.89	0.04
birth order ≥ 5	12.84	-0.03	3.26	-0.25	-6.15	0.18	-8.78	0.35	-12.80	-2.26
<i>Global IQ^b</i>										
1	25.10	-0.15	2.38	-0.06	-12.81	0.73	-15.81	0.70	-16.54	-1.50
2	24.61	-0.06	7.38	-0.04	-14.12	0.59	-19.56	1.01	-22.95	0.76
3	20.16	-0.20	9.86	-0.08	-10.82	0.57	-19.06	0.23	-25.38	1.56
4	12.75	-0.28	11.48	-0.04	-6.19	0.59	-14.88	0.33	-25.97	1.65
6	-14.77	-0.32	-1.55	-0.22	11.98	-0.26	12.57	0.02	3.83	-0.09
7	-21.50	0.75	-11.83	-0.04	12.95	-0.14	22.60	0.08	26.99	-0.50
8	-21.17	1.04	-17.25	0.97	8.19	-0.04	21.94	0.07	43.00	-0.50
9	-17.41	4.07	-16.83	1.78	1.23	0.20	12.09	-0.14	50.72	-0.41
missing	4.04	-2.40	2.09	-1.12	-4.64	-1.14	-3.80	-1.50	-3.48	-2.17
<i>Psychological assessment^b</i>										
1	13.42	0.84	2.90	0.34	-7.32	1.14	-10.38	2.09	-10.52	3.67
2	15.82	1.22	4.44	0.15	-9.07	0.37	-13.19	0.55	-13.52	0.27
3	13.22	0.70	4.58	0.14	-6.06	0.23	-12.05	0.35	-13.54	0.08
4	6.54	0.33	4.63	-0.12	-2.16	-0.06	-7.46	-0.90	-11.39	0.66
6	-8.89	-0.43	-1.50	-0.13	6.29	-0.01	7.74	0.15	4.16	0.21
7	-14.19	-0.53	-7.07	0.21	7.81	-0.28	13.73	0.39	17.72	-1.01
8	-14.54	0.72	-9.80	0.90	6.08	-0.15	14.73	0.14	24.86	-0.65
9	-10.84	1.22	-8.90	1.55	2.86	0.11	10.30	0.21	24.00	-0.37
missing	4.67	-2.16	2.24	-1.08	-4.90	-1.12	-4.36	-1.16	-4.00	-2.04
<i>birth year</i>										
1951	3.92	0.91	-4.04	0.56	5.64	0.38	-1.18	0.78	0.45	1.34
1952	6.66	1.61	-4.13	-0.01	0.23	0.10	-0.37	0.21	0.31	-0.36
1953	8.21	1.30	-4.20	0.56	-1.08	0.18	-0.36	0.84	-0.69	-0.24
1954	-2.08	-0.89	1.67	-0.26	-2.17	-0.28	0.65	-0.01	0.18	0.39
1955	-1.99	-0.45	2.17	-0.38	-2.78	-0.21	-0.29	-0.37	0.21	-0.44
1956	-2.25	-0.85	2.40	-0.12	-3.07	-0.07	-0.14	-0.23	0.14	0.46
1957	-3.73	-0.75	3.10	-0.38	-3.11	-0.04	1.06	0.17	-0.40	-0.56
1958	-5.49	-0.15	3.98	0.27	-3.01	0.13	1.03	-0.17	-0.03	0.99
1959	-7.94	0.76	4.93	-0.68	-1.64	-0.97	0.92	-1.22	0.24	-1.53

^a (1) primary education; (2) Some Secondary education; (3) Full Secondary education 3 years; (4) post-secondary education; (5) Higher

^b Running from low to high.

Table B.4: Impact of hospitalisation and education on mortality hazard IPW, adjacent educations

	Education level ^a				
	(1)	(2)	(3)	(4)	(5)
	<i>base category</i>				
education	—	—	—	—	—
in hospital	4.277** (0.059)	4.374** (0.064)	4.536** (0.127)	4.722** (0.134)	—
hospital experience	1.721** (0.123)	2.748** (0.154)	3.074** (0.271)	3.434** (0.305)	—
	<i>second category</i>				
education	—	-0.464** (0.016)	-0.293** (0.025)	-0.338** (0.034)	-0.123** (0.036)
in hospital	—	4.131** (0.055)	4.342** (0.090)	4.326** (0.132)	4.650** (0.140)
hospital experience	—	1.594** (0.122)	2.627** (0.158)	2.797** (0.272)	3.259** (0.306)

^a (1) primary education; (2) Some Secondary education; (3) Full secondary education; (4) Post-secondary education; (5) University or PhD.

Estimating separate models for (1) and (2); (2) and (3); (3) and (4); (4) and (5); ⁺ $p < 0.05$, ^{**} $p < 0.01$.

Table B.5: Sensitivity: Mortality, (first) hospitalisation and discharge effect with IQ defining the distribution of U given education $e = 1, \dots, 5$ mortality $\delta_M = 0, 1$ and hospitalisation experience $\delta_H = 0, 1$

	Educational levels ^a				
	(1)	(2)	(3)	(4)	(5)
	Mortality effect, $\omega_M(e)$				
<i>Global IQ</i> ^b					
1	0.223**	0.449**	0.474	-0.046	-0.565
2	-0.095 ⁺	0.258**	0.276 ⁺	-0.056	-0.160
3	-0.118**	0.113**	0.143	0.105	0.351
4	-0.148**	-0.013	0.039	-0.051	0.085
6	-0.219**	-0.212**	-0.048	-0.102	0.023
7	-0.066	-0.199**	-0.088	0.037	-0.013
8	0.058	-0.209 ⁺	-0.037	0.030	-0.126
9	0.309	-0.264	0.096	-0.117	-0.145
missing	0.546**	0.171**	0.156 ⁺	-0.017	0.105
	First hospitalisation effect, $\omega_H(e)$				
<i>Global IQ</i> ^b					
1	0.290**	0.375**	0.328**	0.379	0.086
2	0.154**	0.213**	0.317**	0.398**	0.073
3	0.103**	0.219**	0.201**	0.188**	0.207 ⁺
4	0.085**	0.123**	0.078**	0.149**	0.158 ⁺
6	-0.057	0.025	0.105	0.142**	0.177**
7	-0.073	-0.009	0.069**	0.046	0.145**
8	-0.228**	-0.005	-0.048	0.009	-0.003
9	-0.361 ⁺	-0.091	-0.002	0.052	0.031
missing	-0.526**	-1.190**	-1.183**	-1.372**	-1.410**
	Discharge effect, $\omega_D(e)$				
<i>Global IQ</i> ^b					
1	-0.009	-0.028	-0.043	-0.007	0.076
2	0.012	0.000	-0.015	0.032	0.034
3	0.009	0.006	-0.010	0.009	0.006
4	0.005	0.016	-0.001	0.010	-0.006
6	0.020	0.027	-0.001	0.008	0.003
7	0.027	0.020	-0.004	-0.003	0.003
8	0.016	0.011	-0.006	-0.003	0.005
9	0.004	0.040	0.016	0.007	0.016
missing	-0.098	-0.089	-0.118	-0.064	-0.079

^a (1) primary education; (2) Some Secondary education; (3) Full Secondary education 3 years; (4) post-secondary education; (5) Higher

^b Running from low to high.

Based on adding discrete U to propensity score with probabilities of U from observed probabilities for each covariate. No effect would give $\omega_M(e) = 0, \omega_H(e) = 0$ and $\omega_D(e) = 0$. ⁺ $p < 0.05$ and ^{**} $p < 0.01$

Table B.5: Sensitivity: re-admittance and selection effect with IQ defining distribution of U given education $e = 1, \dots, 5$ mortality $\delta_M = 0, 1$ and hospitalisation experience $\delta_H = 0, 1$

	Educational levels ^a				
	(1)	(2)	(3)	(4)	(5)
	Re-admittance effect, $\omega_R(e)$				
<i>Global IQ</i> ^b					
1	0.013	0.039	0.036	-0.062	-0.110
2	-0.018	0.011	-0.015	0.003	0.013
3	-0.017	0.003	-0.009	0.001	0.008
4	-0.010	-0.007	-0.010	-0.014	0.009
6	-0.023	-0.023	-0.009	-0.010	-0.000
7	-0.020	-0.005	-0.006	0.006	-0.001
8	-0.025	0.004	-0.012	0.007	0.002
9	-0.054	-0.049	-0.052	-0.004	-0.004
missing	0.117**	0.109 ⁺	0.117	0.087	0.071
	Selection effect, $\xi(e)$				
<i>Global IQ</i> ^b					
1		-0.765**	-1.238**	-0.793**	-0.368**
2		-0.536**	-0.992**	-0.636**	-0.574**
3		-0.362**	-0.719**	-0.499**	-0.437**
4		-0.181**	-0.504**	-0.364**	-0.389**
6		0.312**	0.172**	-0.051**	-0.129**
7		0.531**	0.561**	0.155**	0.055**
8		0.724**	0.958**	0.340**	0.255**
9		0.928**	1.405**	0.537**	0.528**
missing		-0.082**	-0.119**	0.015	0.002

^a (1) primary education; (2) Some Secondary education; (3) Full Secondary education 3 years; (4) post-secondary education; (5) Higher

^b Running from low to high.

Based on adding discrete U to propensity score with probabilities of U from observed probabilities for each covariate. No effect would give $\omega_R(e) = 0$ and $\xi(e) = 0$. ⁺ $p < 0.05$ and ^{**} $p < 0.01$

Table B.6: Sensitivity: Mortality, (first) hospitalisation and discharge effect with psychological assessment defining distribution of U given education $e = 1, \dots, 5$ mortality $\delta_M = 0, 1$ and hospitalisation experience $\delta_H = 0, 1$

	Educational levels ^a				
	(1)	(2)	(3)	(4)	(5)
	Mortality effect, $\omega_M(e)$				
<i>psychological assessment</i> ^b					
1	0.790**	0.909**	0.757**	0.709**	0.689**
2	0.222**	0.462**	0.561**	0.625**	0.272
3	-0.024	0.243**	0.173 ⁺	0.241 ⁺	0.170
4	-0.283**	-0.100**	0.030	-0.005	0.137
6	-0.403**	-0.387**	-0.161 ⁺	-0.072	-0.104
7	-0.398**	-0.298**	-0.242**	-0.125	-0.111
8	-0.314**	-0.325**	-0.249 ⁺	-0.131	-0.243 ⁺
9	-0.198	-0.264	-0.055	-0.270	-0.198
missing	0.554**	0.228**	0.191**	0.007	0.122
	First hospitalisation effect, $\omega_H(e)$				
<i>psychological assessment</i> ^b					
1	0.306**	0.398**	0.230 ⁺	0.152	0.325 ⁺
2	0.223**	0.321**	0.245**	0.108	0.257**
3	0.135**	0.208**	0.170**	0.181**	0.167**
4	0.022	0.091**	0.141**	0.109**	0.127**
6	-0.016	0.021	0.049	0.081 ⁺	0.091
7	-0.032	0.038	0.046	0.026	0.025
8	0.052	0.018	0.082	0.086	0.085
9	-0.244	0.125	0.033	0.168**	0.138 ⁺
missing	-0.460**	-1.044**	-1.060**	-1.241**	-1.278**
	Discharge effect, $\omega_D(e)$				
<i>psychological assessment</i> ^b					
1	-0.057	-0.049	-0.023	0.008	-0.005
2	-0.010	-0.015	-0.045	-0.020	0.013
3	0.003	0.003	-0.009	0.011	0.010
4	0.020	0.016	-0.007	0.011	0.007
6	0.018	0.026	0.006	0.000	0.004
7	0.038	0.032	0.018	0.009	0.006
8	0.030	0.022	0.016	0.007	0.014
9	0.044	0.022	-0.009	0.032	0.014
missing	-0.093 ⁺	-0.089	-0.107	-0.051	-0.078

^a (1) primary education; (2) Some Secondary education; (3) Full Secondary education 3 years; (4) post-secondary education; (5) Higher

^b Running from low to high.

Based on adding discrete U to propensity score with probabilities of U from observed probabilities for each covariate. No effect would give $\omega_M(e) = 0$, $\omega_H(e) = 0$ and $\omega_D(e) = 0$. ⁺ $p < 0.05$ and ^{**} $p < 0.01$

Table B.6: Sensitivity: re-admittance and selection effect with psychological assessment defining the distribution of U given education $e = 1, \dots, 5$ mortality $\delta_M = 0, 1$ and hospitalisation experience $\delta_H = 0, 1$

	Educational levels ^a				
	(1)	(2)	(3)	(4)	(5)
	Re-admittance effect, $\omega_R(e)$				
<i>psychological assessment</i> ^b					
1	0.068	0.067	0.023	0.036	0.008
2	0.012	0.031	0.034	0.036	-0.005
3	-0.008	0.006	-0.006	-0.010	0.002
4	-0.027	-0.009	0.001	-0.013	0.003
6	-0.024	-0.021	-0.017	0.001	0.003
7	-0.045	-0.021	-0.021	-0.009	0.002
8	-0.034	-0.013	-0.033	-0.008	0.000
9	-0.060	-0.010	-0.006	-0.031	-0.001
missing	0.111**	0.107 ⁺	0.108	0.074	0.070
	Selection effect, $\xi(e)$				
<i>psychological assessment</i> ^b					
1		-0.443**	-0.607**	-0.251**	-0.019
2		-0.364**	-0.496**	-0.185**	-0.018
3		-0.249**	-0.328**	-0.180**	-0.048 ⁺
4		-0.103**	-0.202**	-0.141**	-0.088**
6		0.177**	0.121**	-0.004	-0.056**
7		0.314**	0.349**	0.115**	0.056**
8		0.457**	0.574**	0.230**	0.152**
9		0.606**	0.832**	0.347**	0.273**
missing		-0.092**	-0.127**	0.010	0.004

^a (1) primary education; (2) Some Secondary education; (3) Full Secondary education 3 years; (4) post-secondary education; (5) Higher

^b Running from low to high.

Based on adding discrete U to propensity score with probabilities of U from observed probabilities for each covariate. No effect would give $\omega_R(e) = 0$ and $\xi(e) = 0$. ⁺ $p < 0.05$ and ** $p < 0.01$

Table B.7: Sensitivity: Mortality and (first) hospitalisation effect with mother’s ses or father’s education level defining the distribution of U given education $e = 1, \dots, 5$ mortality $\delta_M = 0, 1$ and hospitalisation experience $\delta_H = 0, 1$

	Educational levels ^a				
	(1)	(2)	(3)	(4)	(5)
<i>Mother’s SES</i>					
	Mortality effect, $\omega_M(e)$				
Unskilled workers	0.211**	0.213**	0.181	0.037	-0.053
Farmers	0.188**	0.235**	0.122	0.359**	0.221
Non-manual (low)	-0.369**	-0.255**	-0.139	-0.041	-0.084
Non-manual (medium)	0.137**	-0.050	-0.031	-0.036	-0.070
Non-manual (high)	0.181	0.101	0.146	-0.104	0.006
not classified	0.258 ⁺	-0.052	0.008	-0.178	-0.064
missing	1.046**	0.624**	0.460 ⁺	0.079	0.144
<i>Father’s education</i>					
less than 9 years	-0.388**	-0.131**	-0.094	-0.039	0.027
9-10 years	-0.120	-0.104	0.133	0.058	0.123
Full secondary	0.004	-0.094	-0.111	-0.019	-0.073
University < 3 years	0.178	0.040	0.036	-0.227	-0.055
University \geq 3 years	0.218 ⁺	-0.032	-0.067	-0.206	-0.191 ⁺
PhD studies	0.894**	0.083	0.138	0.115	-0.294
missing	0.643**	0.423**	0.422**	0.281**	0.438**
<i>Mother’s SES</i>					
	First hospitalisation effect, $\omega_H(e)$				
Unskilled workers	0.008	-0.000	-0.104	-0.074	-0.017
Farmers	0.040	0.070	0.052	0.130 ⁺	0.187**
Non-manual (low)	-0.031	-0.062**	-0.005	0.015	-0.024
Non-manual (medium)	-0.044	-0.020**	-0.060	-0.045	-0.070 ⁺
Non-manual (high)	-0.057	-0.077	-0.160 ⁺	-0.175**	-0.115 ⁺
not classified	-0.011	0.023	-0.110	-0.053	-0.044
missing	-0.355**	-0.129	-0.233	-0.216	-0.123
<i>Father’s education</i>					
less than 9 years	0.099**	0.069**	0.128**	0.119**	0.130**
9-10 years	-0.031	-0.021	0.001	0.023	-0.019
Full secondary	-0.081	-0.054	-0.055	-0.085 ⁺	0.032
University < 3 years	-0.032	-0.074	-0.191 ⁺	-0.112	-0.094
University \geq 3 years	-0.325**	-0.166 ⁺	-0.238**	-0.225**	-0.176**
PhD studies	-0.052	0.009	-0.312	-0.070	-0.298**
missing	-0.075**	-0.023	-0.020	0.020	0.059

^a (1) primary education; (2) Some Secondary education; (3) Full Secondary education 3 years; (4) post-secondary education; (5) Higher

Based on adding discrete U to propensity score with probabilities of U from observed probabilities for each covariate. No effect would give $\omega_M(e) = 0$ and $\omega_H(e) = 0$.

⁺ $p < 0.05$ and ** $p < 0.01$

Table B.7: Sensitivity: discharge and re-admittance effect with mother’s ses or father’s education variable defining the distribution of U given education $e = 1, \dots, 5$ mortality $\delta_M = 0, 1$ and hospitalisation experience $\delta_H = 0, 1$

	Educational levels ^a				
	(1)	(2)	(3)	(4)	(5)
<i>Mother’s SES</i>					
	Discharge effect, $\omega_D(e)$				
Unskilled workers	-0.006	-0.002	-0.031	0.010	0.007
Farmers	-0.010	-0.005	-0.030	-0.011	0.010
Non-manual (low)	0.013	0.021	-0.009	0.011	0.016
Non-manual (medium)	-0.012	0.009	0.002	-0.011	0.005
Non-manual (high)	0.010	0.004	0.004	0.010	0.010
not classified	-0.018	0.021	-0.002	0.018	0.004
missing	0.010	-0.002	0.056	0.008	0.028
<i>Father’s education</i>					
less than 9 years	0.006	0.007	-0.001	0.001	-0.003
9-10 years	0.003	0.007	-0.027	0.009	-0.021
Full secondary	0.002	0.016	-0.004	-0.003	0.009
University < 3 years	0.011	0.007	-0.003	0.032	0.023
University \geq 3 years	-0.013	0.008	0.011	0.002	0.019
PhD studies	-0.050	0.004	0.018	-0.001	0.033
missing	-0.014	-0.006	-0.017	-0.010	-0.011
	Re-admittance effect, $\omega_R(e)$				
<i>Mother’s SES</i>					
Unskilled workers	0.018	0.013	0.012	-0.002	-0.004
Farmers	0.014	0.019	0.017	0.014	0.005
Non-manual (low)	-0.020	-0.012	0.004	-0.011	0.004
Non-manual (medium)	0.011	-0.001	-0.014	0.006	0.002
Non-manual (high)	-0.006	0.012	-0.020	-0.003	0.008
not classified	0.027	0.000	-0.008	-0.011	0.005
missing	0.003	0.011	-0.050	0.030	-0.002
<i>Father’s education</i>					
less than 9 years	-0.012	-0.003	0.003	-0.001	0.004
9-10 years	-0.009	-0.001	0.016	-0.007	0.023
Full secondary	0.001	0.001	-0.008	0.003	0.003
University < 3 years	-0.004	0.004	-0.010	-0.033	-0.006
University \geq 3 years	0.020	0.014	-0.035	-0.002	-0.004
PhD studies	0.074	-0.049	-0.031	-0.003	-0.016
missing	0.014	0.015	0.006	0.011	0.016

^a (1) primary education; (2) Some Secondary education; (3) Full Secondary education 3 years; (4) post-secondary education; (5) Higher

Based on adding discrete U to propensity score with probabilities of U from observed probabilities for each covariate. No effect would give $\omega_D(e) = 0$ and $\omega_R(e) = 0$.

⁺ $p < 0.05$ and ^{**} $p < 0.01$

Table B.7: Sensitivity: selection effect with mother’s ses or father’s education variable defining the distribution of U given education $e = 1, \dots, 5$ mortality $\delta_M = 0, 1$ and hospitalisation experience $\delta_H = 0, 1$

	Educational levels ^a				
	(1)	(2)	(3)	(4)	(5)
	Selection effect, $\xi(e)$				
<i>Mother’s SES</i>					
Unskilled workers		-0.107**	-0.249**	-0.079**	-0.048
Farmers		-0.125**	-0.231**	-0.170**	-0.159**
Non-manual (low)		-0.220**	-0.144**	-0.147**	-0.200**
Non-manual (medium)		0.406**	0.492**	0.221**	0.220**
Non-manual (high)		0.467**	0.586**	0.309**	0.305**
not classified		0.459**	0.746**	0.258**	0.452**
missing		-0.089**	0.003	-0.016	0.061
<i>Father’s education</i>					
less than 9 years		-0.319**	-0.463**	-0.272**	-0.360**
9-10 years		0.131**	0.123**	-0.010	-0.030
Full secondary		0.409**	0.389**	0.177**	0.044**
University < 3 years		0.514**	0.477**	0.239**	0.116**
University \geq 3 years		0.723**	1.010**	0.479**	0.620**
PhD studies		0.813**	1.298**	0.568**	0.885**
missing		-0.167**	-0.084**	-0.080**	0.002

^a (1) primary education; (2) Some Secondary education; (3) Full Secondary education 3 years; (4) post-secondary education; (5) Higher

Based on adding discrete U to propensity score with probabilities of U from observed probabilities for each covariate. No effect would give $\xi(e) = 0$. ⁺ $p < 0.05$ and ^{**} $p < 0.01$

Table B.8: Sensitivity analysis: IPW-ToE model with psychological assessment, mother’s ses or father’s education defining the distribution of U given education $e = 1, \dots, 5$ mortality $\delta_M = 0, 1$ and hospitalisation experience $\delta_H = 0, 1$

	Educational levels ^a				
	(1)	(2)	(3)	(4)	(5)
	Education effect				
original	–	–0.585**	–0.863**	–1.216**	–1.303**
<i>psychological assessment</i> ^b					
1		–0.566	–0.819	–1.171	–1.259
2		–0.580	–0.835	–1.178	–1.276
3		–0.588	–0.857	–1.202	–1.291
4		–0.588	–0.868	–1.220	–1.301
6		–0.574	–0.848	–1.200	–1.288
7		–0.576	–0.844	–1.196	–1.281
8		–0.583	–0.850	–1.200	–1.276
9		–0.586	–0.860	–1.210	–1.293
missing		–0.576	–0.849	–1.204	–1.290
<i>Mother’s SES</i>					
Unskilled workers		–0.585	–0.859	–1.213	–1.300
Farmers		–0.584	–0.859	–1.205	–1.288
Non-manual (low)		–0.599	–0.879	–1.230	–1.320
Non-manual (medium)		–0.598	–0.875	–1.220	–1.306
Non-manual (high)		–0.587	–0.870	–1.219	–1.309
not classified		–0.587	–0.865	–1.216	–1.301
missing		–0.580	–0.858	–1.211	–1.298
<i>Father’s education</i>					
less than 9 years		–0.625	–0.923	–1.273	–1.364
9-10 years		–0.584	–0.863	–1.215	–1.302
Full secondary		–0.586	–0.861	–1.214	–1.297
University < 3 years		–0.588	–0.868	–1.218	–1.309
University ≥ 3 years		–0.591	–0.870	–1.218	–1.290
PhD studies		–0.586	–0.868	–1.221	–1.304
missing		–0.555	–0.829	–1.180	–1.263

^a (1) primary education; (2) Some Secondary education; (3) Full Secondary education 3 years; (4) post-secondary education; (5) Higher

^b Running from low to high.

Based on adding discrete U to propensity score with probabilities of U from observed probabilities for each covariate. Significance of difference with original estimates in Table 2, ⁺ $p < 0.05$ and ^{**} $p < 0.01$

Table B.8: Sensitivity analysis IPW-ToE model (continued)

	Educational levels ^a				
	(1)	(2)	(3)	(4)	(5)
	in hospital effect				
original	4.221**	4.038**	3.966**	3.718**	3.660**
<i>psychological assessment</i> ^b					
1	4.223	4.053	3.991	3.735	3.678
2	4.226	4.047	3.994	3.792	3.662
3	4.225	4.040	3.969	3.724	3.660
4	4.222	4.037	3.963	3.706	3.651
6	4.224	4.046	3.979	3.716	3.673
7	4.217	4.044	3.984	3.719	3.686
8	4.223	4.044	3.982	3.721	3.684
9	4.218	4.037	3.966	3.715	3.657
missing	4.214	4.039	3.979	3.731	3.670
<i>Mother's SES</i>					
Unskilled workers	4.222	4.039	3.969	3.729	3.667
Farmers	4.222	4.040	3.972	3.733	3.666
Non-manual (low)	4.216	4.027	3.955	3.705	3.640
Non-manual (medium)	4.224	4.024	3.967	3.665	3.719
Non-manual (high)	4.221	4.039	3.966	3.704	3.660
not classified	4.225	4.036	3.965	3.709	3.652
missing	4.225	4.042	3.969	3.722	3.663
<i>Father's education</i>					
less than 9 years	4.201	4.010	3.941	3.703	3.552
9-10 years	4.222	4.039	3.965	3.714	3.658
Full secondary	4.224	4.039	3.970	3.697	3.679
University < 3 years	4.221	4.039	3.965	3.713	3.664
University ≥ 3 years	4.226	4.042	3.969	3.703	3.689
PhD studies	4.220	4.034	3.962	3.715	3.666
missing	4.239	4.060	3.987	3.748	3.683

^a (1) primary education; (2) Some Secondary education; (3) Full Secondary education 3 years; (4) post-secondary education; (5) Higher

^b Running from low to high.

Based on adding discrete U to propensity score with probabilities of U from observed probabilities for each covariate. Significance of difference with original estimates in Table 2,

⁺ $p < 0.05$ and ^{**} $p < 0.01$

Table B.8: Sensitivity analysis IPW-ToE model (continued)

	Educational levels ^a				
	(1)	(2)	(3)	(4)	(5)
	Hospital experience effect				
original	1.546**	1.370**	1.236**	0.837**	0.730**
<i>psychological assessment</i> ^b					
1	1.547	1.389	1.273	0.867	0.765
2	1.547	1.375	1.268	0.861	0.736
3	1.547	1.367	1.234	0.825	0.728
4	1.546	1.366	1.231	0.825	0.725
6	1.555	1.386	1.258	0.854	0.748
7	1.542	1.381	1.262	0.862	0.751
8	1.551	1.386	1.260	0.860	0.756
9	1.542	1.369	1.236	0.847	0.735
missing	1.545	1.378	1.262	0.858	0.754
<i>Mother's SES</i>					
Unskilled workers	1.545	1.369	1.238	0.835	0.731
Farmers	1.546	1.370	1.241	0.844	0.731
Non-manual (low)	1.535	1.351	1.217	0.813	0.708
Non-manual (medium)	1.546	1.352	1.232	0.805	0.725
Non-manual (high)	1.546	1.372	1.235	0.838	0.725
not classified	1.559	1.375	1.240	0.844	0.728
missing	1.554	1.377	1.243	0.845	0.737
<i>Father's education</i>					
less than 9 years	1.494	1.313	1.169	0.763	0.685
9-10 years	1.548	1.372	1.235	0.840	0.730
Full secondary	1.547	1.371	1.238	0.835	0.736
University < 3 years	1.537	1.362	1.225	0.837	0.727
University ≥ 3 years	1.527	1.359	1.215	0.817	0.732
PhD studies	1.538	1.356	1.222	0.825	0.721
missing	1.588	1.416	1.283	0.886	0.783

^a (1) primary education; (2) Some Secondary education; (3) Full Secondary education 3 years; (4) post-secondary education; (5) Higher

^b Running from low to high.

Based on adding discrete U to propensity score with probabilities of U from observed probabilities for each covariate. Significance of difference with original estimates in Table 2,

⁺ $p < 0.05$ and ^{**} $p < 0.01$

Appendix C Full tables with parameter estimates

Table C.1: Parameters of (first) CVD hospitalisation hazard (Gompertz)

	Model ^a		
	ToE	Toe IPW	Toe IPW rob
primary education	—	—	—
Some Secondary education	−0.084** (0.009)	−0.043** (0.009)	−0.040** (0.009)
Full Secondary education	−0.248** (0.012)	−0.150** (0.012)	−0.153** (0.012)
Post-secondary education	−0.308** (0.012)	−0.178** (0.012)	−0.180** (0.012)
University or PhD	−0.457** (0.012)	−0.311** (0.012)	−0.318** (0.012)
γ	0.099** (0.000)	0.098** (0.000)	0.095** (0.000)
v_1	−9.517** (0.025)	−9.511** (0.025)	−9.410** (0.032)
v_2	−9.436** (0.084)	−9.571** (0.094)	−9.429** (0.103)

^a ToE: Timing-of-events; ToE IPW; Timing-of-events with Inverse propensity weighting; ToE IPW rob; Robust Timing-of-events with Inverse propensity weighting.

⁺ $p < 0.05$, ** $p < 0.01$.

Table C.2: Parameters of CVD hospitalisation discharge hazard (Weibull)

	Model ^a		
	ToE	Toe IPW	Toe IPW rob
primary education	—	—	—
Some Secondary education	0.078** (0.007)	0.035** (0.007)	0.046** (0.007)
Full Secondary education	0.089** (0.010)	0.026** (0.010)	0.026** (0.010)
Post-secondary education	0.149** (0.010)	0.031** (0.010)	0.065** (0.010)
University or PhD	0.206** (0.010)	0.096** (0.010)	0.108** (0.010)
α	0.991 (0.010)	0.966** (0.002)	0.975** (0.002)
v_1	0.753** (0.006)	0.765** (0.006)	0.671** (0.016)
v_2	-0.608** (0.008)	-0.631** (0.009)	-0.742** (0.018)

^a ToE: Timing-of-events; ToE IPW; Timing-of-events with Inverse propensity weighting; ToE IPW rob; Robust Timing-of-events with Inverse propensity weighting.

⁺ $p < 0.05$, ** $p < 0.01$.

Table C.3: Parameters of CVD re-hospitalisation hazard (Weibull)

	Model ^a		
	ToE	Toe IPW	Toe IPW rob
primary education	–	–	–
Some Secondary education	–0.024** (0.007)	0.004 (0.007)	0.004 (0.007)
Full Secondary education	–0.117** (0.010)	–0.073** (0.010)	–0.069** (0.010)
Post-secondary education	–0.205** (0.010)	–0.137** (0.010)	–0.142** (0.010)
University or PhD	–0.254** (0.010)	–0.154** (0.010)	–0.155** (0.010)
α	0.348 (0.001)	0.348** (0.001)	0.348** (0.001)
v_1	–0.603** (0.006)	–0.624** (0.006)	–0.571** (0.015)
v_2	0.148** (0.008)	0.130** (0.008)	0.135** (0.016)

^a ToE: Timing-of-events; ToE IPW; Timing-of-events with Inverse propensity weighting; ToE IPW rob; Robust Timing-of-events with Inverse propensity weighting.

⁺ $p < 0.05$, ** $p < 0.01$.

Table C.4: Parameters of mortality hazard

	Model ^a		
	ToE	Toe IPW	Toe IPW rob
<i>primary education</i>	—	—	—
in hospital	4.213** (0.055)	4.221** (0.060)	4.227** (0.057)
hospital experience	1.830** (0.108)	1.546** (0.110)	1.637** (0.108)
<i>Some Secondary education</i>	-0.587** (0.016)	-0.585** (0.016)	-0.597** (0.016)
in hospital	4.016** (0.053)	4.038** (0.055)	4.002** (0.052)
hospital experience	1.620** (0.108)	1.370** (0.109)	1.423** (0.107)
<i>Full Secondary education</i>	-0.928** (0.025)	-0.863** (0.025)	-0.873** (0.025)
in hospital	3.854** (0.085)	3.966** (0.084)	4.023** (0.082)
hospital experience	1.400** (0.114)	1.236** (0.115)	1.321** (0.113)
<i>Post-secondary education</i>	-1.300** (0.028)	-1.216** (0.027)	-1.238** (0.027)
in hospital	3.642** (0.096)	3.718** (0.090)	3.692** (0.089)
hospital experience	1.096** (0.116)	0.837** (0.117)	0.915** (0.115)
<i>University or PhD</i>	-1.443** (0.027)	-1.303** (0.027)	-1.325** (0.027)
in hospital	3.546** (0.106)	3.660** (0.094)	3.614** (0.093)
hospital experience	0.928** (0.117)	0.730** (0.118)	0.796** (0.116)
γ hospital experience	-0.017** (0.002)	-0.010** (0.002)	-0.013** (0.002)
γ constant	0.072** (0.001)	0.067** (0.001)	0.065** (0.001)
v_1	-9.569** (0.045)	-9.398** (0.046)	-9.378** (0.057)
v_2	-7.834** (0.042)	-7.682** (0.043)	-7.841** (0.054)
p_1	0.860** (0.010)	0.859** (0.010)	0.864** (0.012)

^a ToE: Timing-of-events; ToE IPW; Timing-of-events with Inverse propensity weighting; ToE IPW rob; Robust Timing-of-events with Inverse propensity weighting.

⁺ $p < 0.05$, ** $p < 0.01$.

Table C.5: Parameters control variables in Robust Timing-of-events with Inverse propensity weighting

	Hazard ^a			
	Mortality	First	Discharge	Re-admit
<i>SES mother at birth</i>				
not classified	0.015	0.062**	0.063**	-0.020
Unskilled workers	0.110**	0.020	0.041**	0.039**
Farmers	-0.191**	-0.099**	-0.001	-0.046**
non-manual (low)	0.078**	-0.017	-0.012	-0.041**
non-manual (intermediate)	0.144**	-0.091**	0.034 ⁺	-0.029
non-manual (high)	0.129**	-0.055 ⁺	0.031	-0.034
missing	0.738**	-0.275**	0.045	-0.009
<i>father's education</i>				
less than 9 years	-0.055**	0.000	0.003	-0.032**
9-10 years	-0.001	-0.030	-0.005	-0.067**
Full secondary education	0.024	-0.002	-0.016	-0.026 ⁺
university (< 3 years)	0.150**	-0.007	0.013	-0.014
university (≥ 3 years)	0.187**	-0.137**	-0.076**	-0.159**
PhD	0.199**	0.203**	0.296**	0.240**
missing	0.232**	-0.048**	-0.021	-0.002
<i>IQ measurement</i>				
1	0.025	0.146**	-0.233**	0.168**
2	0.030	0.140**	0.043**	0.075**
3	0.090**	0.089**	0.005	-0.032**
4	0.027	0.017	-0.005	0.003
6	-0.020	-0.020	0.009	-0.020 ⁺
7	0.064**	-0.059**	0.072**	-0.064**
8	0.089**	-0.118**	-0.001	-0.038**
9	0.253**	-0.102**	0.049**	0.100**
missing	0.233**	-0.401**	-0.269**	0.188**

^a Mortality: mortality hazard; First: first CVD hospitalisation hazard; discharge: CVD hospitalisation discharge hazard; re-admit: CVD re-hospitalisation hazard.

Reference category: mother skilled worker and secondary education (max 2 years) and IQ level 5. ⁺ $p < 0.05$, ** $p < 0.01$.

Table C.5: Parameters control variables in Robust Timing-of-events with Inverse propensity weighting (continued)

	Hazard ^a			
	Mortality	First	Discharge	Re-admit
<i>Psychological assessment</i>				
1	0.843**	0.151**	-0.062**	0.186**
2	0.460**	0.186**	-0.092**	0.046**
3	0.239**	0.112**	-0.096**	0.074**
4	0.098**	0.043**	-0.033**	-0.009
6	-0.116**	0.014	0.016	-0.014
7	-0.107**	0.018	0.013	-0.026 ⁺
8	-0.059 ⁺	0.059**	0.107**	0.028 ⁺
9	-0.087	0.023	0.005	-0.071**
missing	0.707**	0.326**	-0.065 ⁺	0.163**
<i>birth year</i>				
1951	0.084**	0.012	0.035**	-0.005
1952	0.064**	-0.002	0.045**	-0.026**
1953	0.009	0.042**	0.076**	-0.094**
1954	-0.262**	0.036 ⁺	0.119**	-0.072**
1955	-0.220**	0.074**	0.088**	-0.117**
1956	-0.301**	0.015	0.105**	-0.143**
1957	-0.273**	0.046**	0.125**	-0.156**
1958	-0.315**	0.080**	0.103**	-0.160**
1959	-0.711**	-1.517**	0.219**	-0.086**
<i>birth info</i>				
mother < 20 at birth	0.123**	0.055**	0.023 ⁺	0.034**
father > 40 at birth	0.206**	-0.245**	-0.010	-0.017
family size 1	0.051 ⁺	0.041**	0.003	0.048**
family size 3	-0.020	-0.033**	0.031**	0.017 ⁺
family size 4	0.037	-0.030 ⁺	0.048**	0.038**
family size 5 or higher	0.056 ⁺	-0.051**	-0.014	0.043**
birth order 2	-0.035 ⁺	0.030**	0.038**	0.011
birth order 3	-0.074**	0.112**	0.051**	0.031**
birth order 4	-0.155**	0.128**	0.027 ⁺	0.014
birth order 5 or higher	0.001	0.176**	0.100**	-0.052**

^a Mortality: mortality hazard; First: first CVD hospitalisation hazard; discharge: CVD hospitalisation discharge hazard; re-admit: CVD re-hospitalisation hazard.
Reference category: 1950, psychological assessment 5. ⁺ $p < 0.05$, ** $p < 0.01$.

Table C.6: Parameters education choice, sequential probit model

	Education choice ^a			
	(1)	(2)	(3)	(4)
<i>SES mother at birth</i>				
not classified	0.020	0.009	0.020	0.077**
Unskilled workers	-0.016 ⁺	0.003	-0.004	0.065**
Farmers	-0.120**	0.046**	-0.003	0.039**
non-manual (low)	0.178**	0.288**	0.114**	0.187**
non-manual (intermediate)	0.228**	0.372**	0.187**	0.247**
non-manual (high)	0.230**	0.521**	0.146**	0.369**
missing	0.013	0.130**	0.059 ⁺	0.178**
<i>father's education</i>				
less than 9 years	-0.229**	-0.174**	-0.065**	-0.098**
9-10 years	-0.080**	0.014	-0.006	0.007
Full secondary education	0.083**	0.168**	0.115**	0.033**
university (< 3 years)	0.169**	0.243**	0.165**	0.090**
university (≥ 3 years)	0.303**	0.635**	0.318**	0.462**
PhD	0.409**	0.955**	0.436**	0.775**
missing	-0.195**	-0.101**	-0.008	0.038 ⁺
<i>IQ measurement</i>				
1	-0.637**	-0.848**	-0.526**	-0.151 ⁺
2	-0.463**	-0.657**	-0.400**	-0.311**
3	-0.324**	-0.437**	-0.265**	-0.164**
4	-0.183**	-0.258**	-0.135**	-0.120**
6	0.224**	0.248**	0.117**	0.094**
7	0.441**	0.563**	0.265**	0.216**
8	0.651**	0.921**	0.413**	0.352**
9	0.841**	1.332**	0.585**	0.571**
missing	-0.264**	0.136**	0.166**	0.169**

^a (1) some secondary education over primary education; (2) full secondary education over some secondary; (3) post-secondary over full secondary education; (4) higher over post-secondary education.

⁺ $p < 0.05$, ** $p < 0.01$.

Table C.6: (continued)

	Education choice ^a			
	(1)	(2)	(3)	(4)
<i>Psychological assessment</i>				
1	-0.240**	-0.268**	-0.038	0.173**
2	-0.196**	-0.198**	-0.012	0.139**
3	-0.145**	-0.111**	-0.042**	0.088**
4	-0.081**	-0.063**	-0.037**	0.032**
6	0.074**	0.094**	0.034**	0.025 ⁺
7	0.155**	0.226**	0.097**	0.080**
8	0.233**	0.357**	0.151**	0.119**
9	0.292**	0.489**	0.201**	0.171**
missing	-0.189**	-0.058 ⁺	0.014	0.113 ⁺
<i>birth year</i>				
1951	-0.018 ⁺	-0.065 ⁺	0.081 ⁺	0.022
1952	-0.042**	-0.043**	0.197**	0.019
1953	-0.097**	-0.080**	0.195**	-0.010
1954	0.138**	-0.204**	0.219**	-0.028
1955	0.129**	-0.230**	0.216**	-0.014
1956	0.137**	-0.230**	0.230**	-0.012
1957	0.172**	-0.238**	0.232**	-0.049**
1958	0.210**	-0.259**	0.231**	-0.047**
1959	0.618**	-0.184**	0.203**	-0.092**
<i>birth info</i>				
mother < 20 at birth	-0.164**	-0.204**	-0.172**	-0.093**
father > 40 at birth	0.028**	0.071**	0.058**	0.047**
family size 1	0.014	0.024**	-0.003	0.042
family size 3	-0.066**	-0.061**	-0.006	-0.023
family size 4	-0.140**	-0.135**	-0.034**	-0.053**
family size 5 or higher	-0.214**	-0.243**	-0.067**	-0.059**
birth order 2	-0.050**	-0.066**	-0.029**	-0.008
birth order 3	-0.046**	-0.050**	-0.031**	0.007
birth order 4	-0.024 ⁺	-0.045**	-0.041 ⁺	0.016
birth order 5 or higher	-0.040**	-0.011	-0.034	-0.043
constant	0.925**	0.926**		

^a (1) some secondary education over primary education; (2) full secondary education over some secondary; (3) post-secondary over full secondary education; (4) higher over post-secondary education.

⁺ $p < 0.05$, ** $p < 0.01$.

Table C.7: Parameters of Cause specific mortality hazard (IPW)

	Cause of death ^a				
	IHD	Stroke	other CVD	other natural	external
primary education	–	–	–	–	–
in hospital	4.630** (0.138)	6.671** (0.154)	5.537** (0.135)	4.110** (0.088)	2.547** (0.261)
hospital experience	4.370** (0.287)	3.925** (0.632)	3.618** (0.332)	2.537** (0.155)	0.564** (0.066)
Some Secondary education	–0.363** (0.049)	–0.183 (0.096)	–0.416** (0.070)	–0.460** (0.024)	–0.810** (0.025)
in hospital	4.421** (0.122)	6.405** (0.147)	5.304** (0.123)	3.974** (0.080)	2.028** (0.272)
hospital experience	4.292** (0.285)	3.351** (0.634)	3.685** (0.329)	2.348** (0.154)	0.184** (0.063)
Full Secondary education	–0.535** (0.072)	–0.724** (0.158)	–0.555** (0.102)	–0.734** (0.036)	–1.131** (0.043)
in hospital	3.803** (0.274)	6.148** (0.236)	5.120** (0.211)	4.022** (0.112)	2.685** (0.368)
hospital experience	4.184** (0.296)	3.345** (0.671)	3.312** (0.347)	2.239** (0.161)	0.112 (0.111)
Post-secondary education	–0.832** (0.078)	–1.316** (0.193)	–0.947** (0.114)	–1.065** (0.039)	–1.514** (0.048)
in hospital	3.904** (0.257)	5.914** (0.254)	4.690** (0.249)	3.806** (0.119)	1.872** (0.525)
hospital experience	3.653** (0.303)	2.306** (0.744)	3.013** (0.351)	1.809** (0.164)	–0.030 (0.112)
University or PhD	–1.107** (0.082)	–0.826** (0.150)	–0.939** (0.108)	–1.050** (0.037)	–1.789** (0.051)
in hospital	3.334** (0.333)	5.722** (0.280)	4.673** (0.253)	3.791** (0.120)	1.018 (0.870)
hospital experience	3.954** (0.298)	2.551** (0.721)	2.542** (0.367)	1.750** (0.165)	–0.756** (0.161)
γ hospital experience	–0.059** (0.005)	–0.054** (0.012)	–0.037** (0.006)	–0.025** (0.003)	–
γ constant	0.125** (0.003)	0.079** (0.004)	0.088** (0.003)	0.091** (0.001)	–
v_1	–14.280** (0.136)	–13.424** (0.200)	–13.328** (0.158)	–11.584** (0.073)	–7.237** (0.042)
v_2	–13.343** (0.146)	–12.814** (0.228)	–11.937** (0.162)	–9.423** (0.067)	–6.274** (0.072)
p_1	0.856** (0.014)				

^a IHD: Ischemic Heart Disease
⁺ $p < 0.05$, ^{**} $p < 0.01$.

Appendix D Additional figures

Figure D.1: Propensity score p_1 overlap by education level

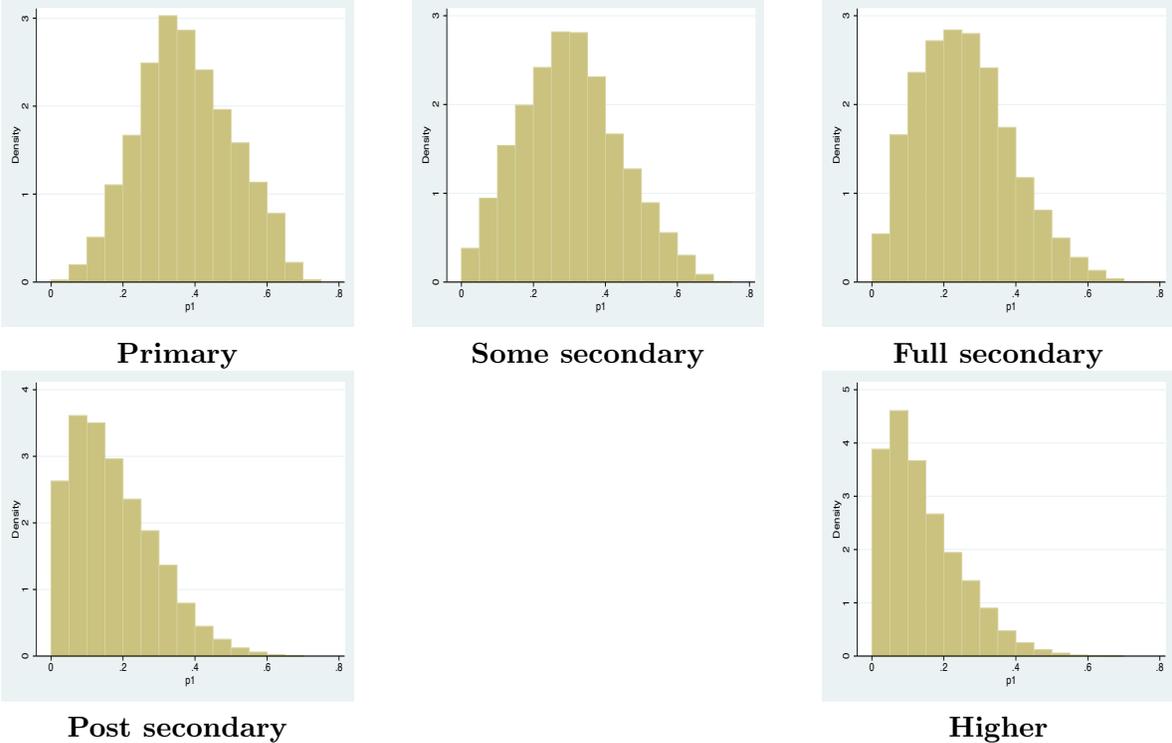


Figure D.2: Propensity score p_2 overlap by education level

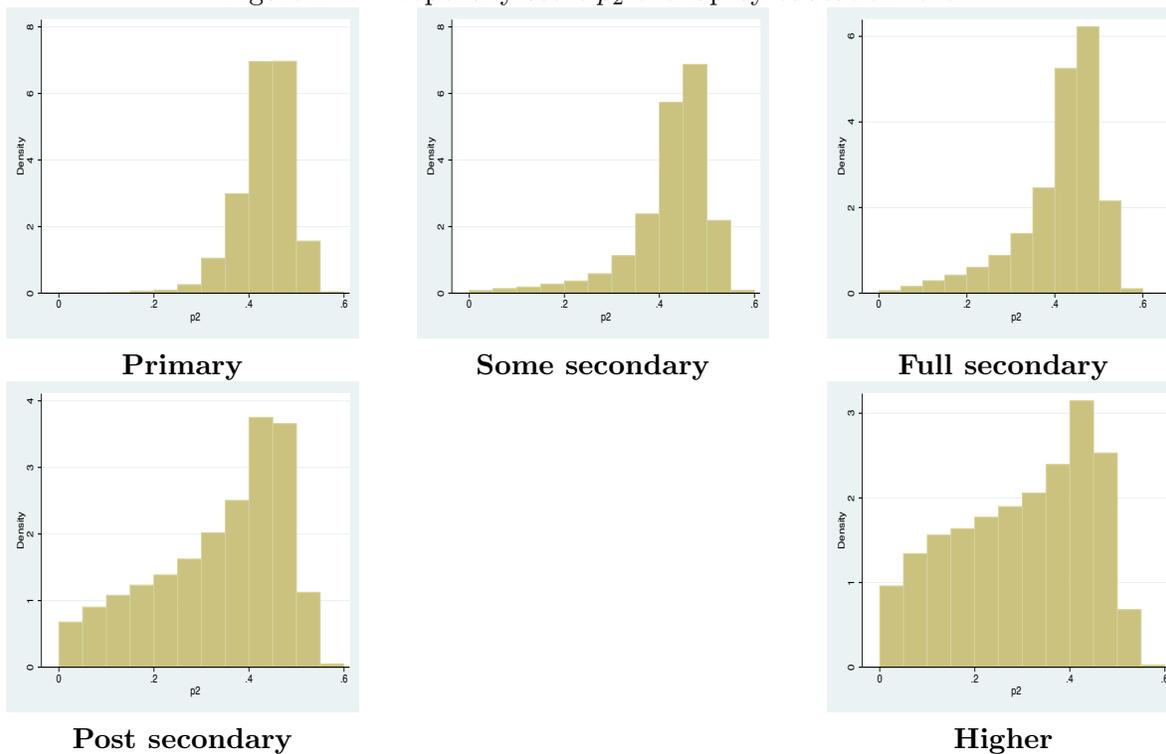


Figure D.3: Propensity score p_3 overlap by education level

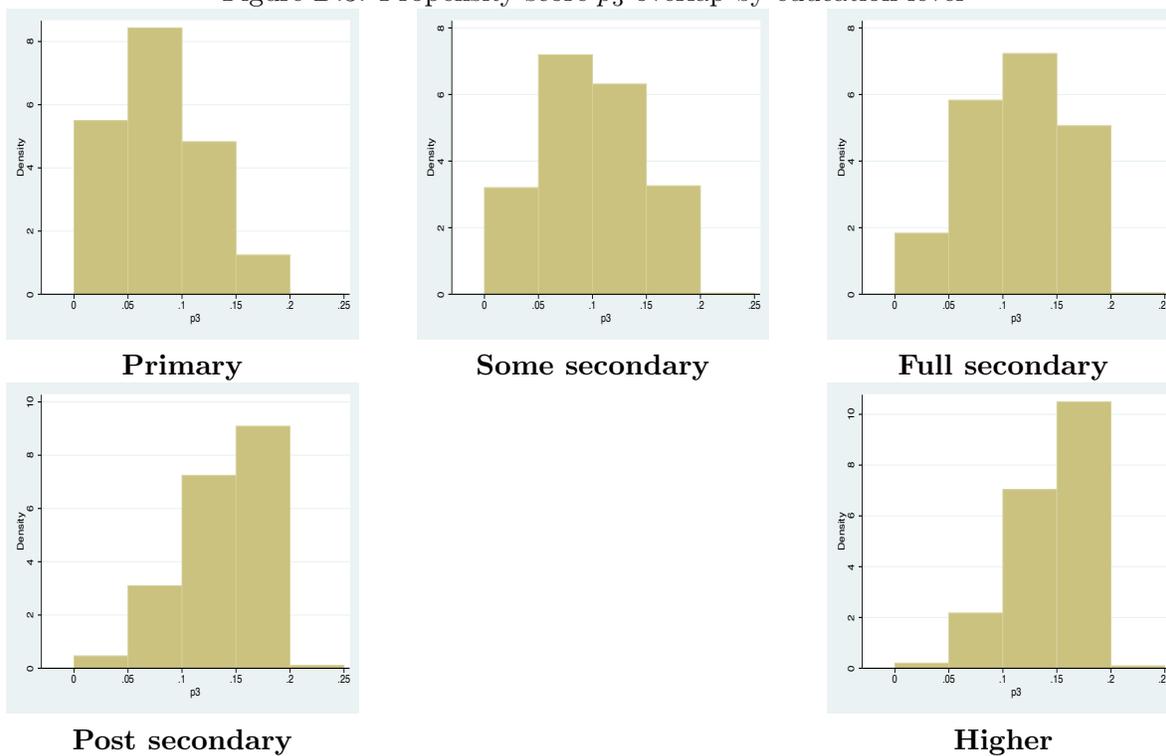


Figure D.4: Propensity score p_4 overlap by education level

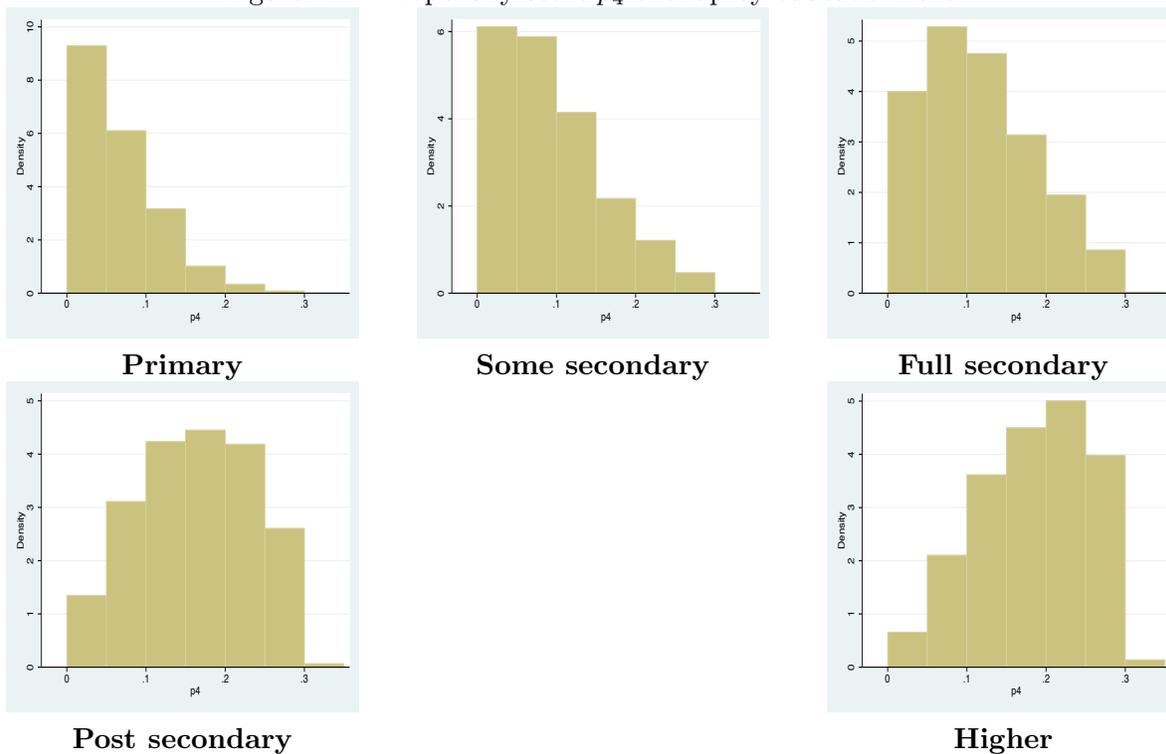


Figure D.5: Propensity score p_5 overlap by education level

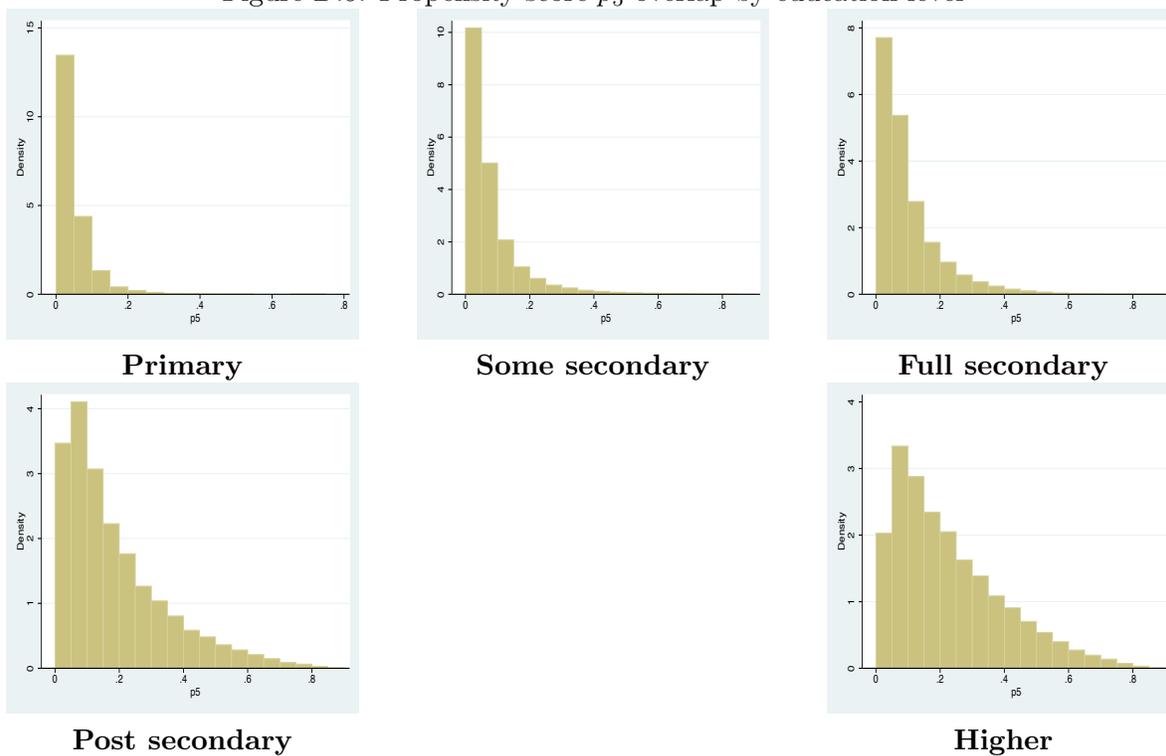


Figure D.6: Cumulative incidence curves by cause of death (CVD), hospitalisation and education level

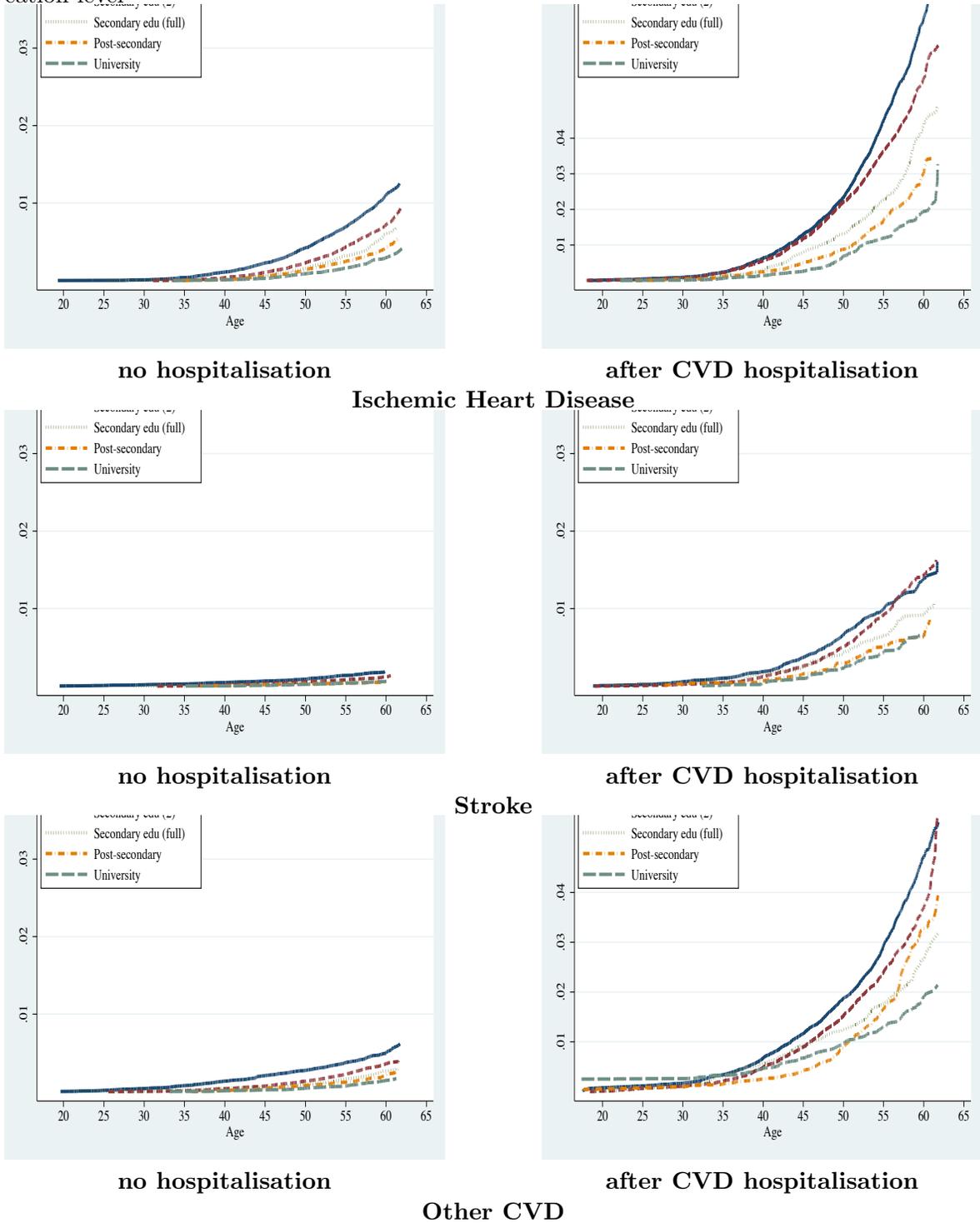


Figure D.7: Cumulative incidence curves by cause of death, hospitalisation and education level, external or other causes

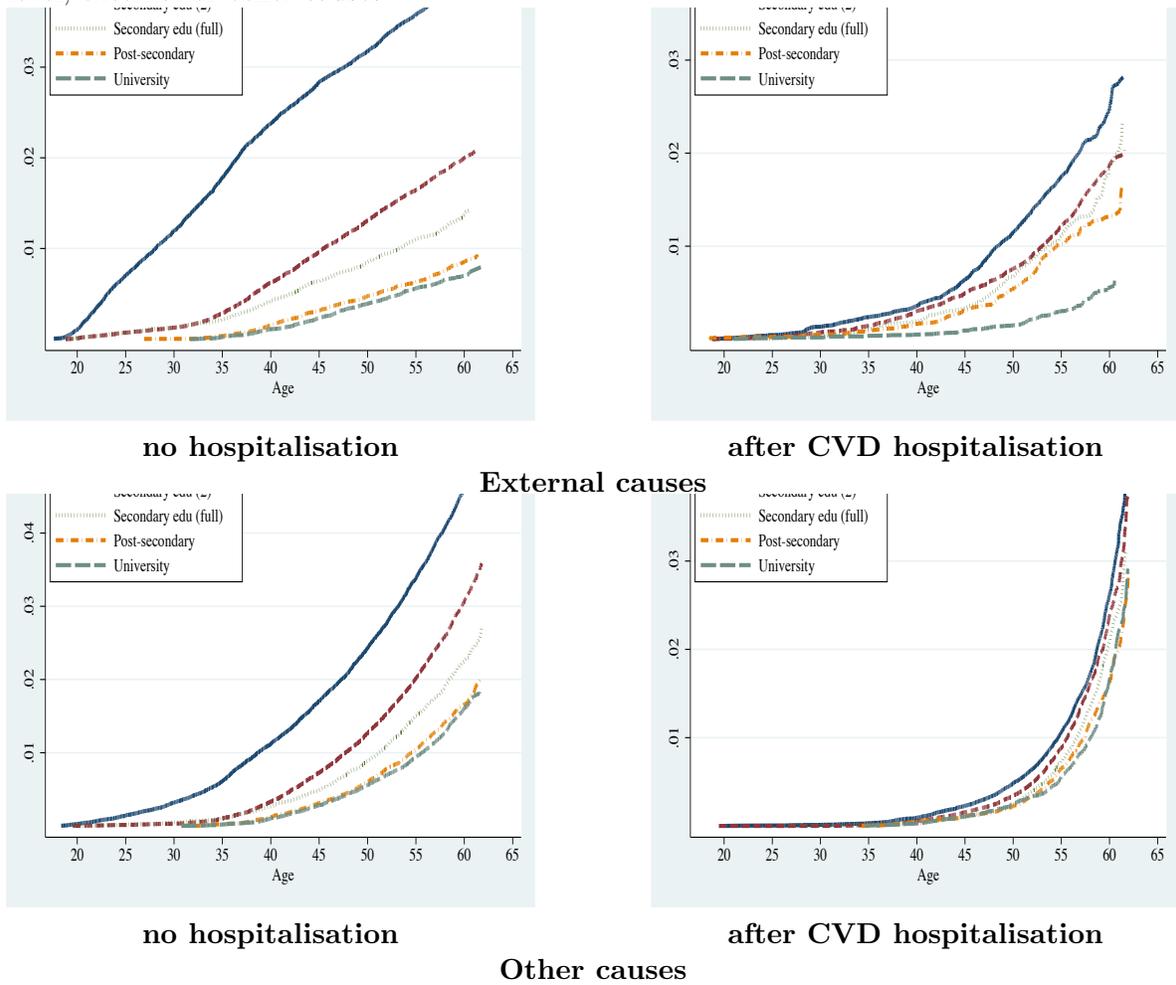


Figure D.8: Simulated Cumulative incidence curves by cause of death (CVD) and education level

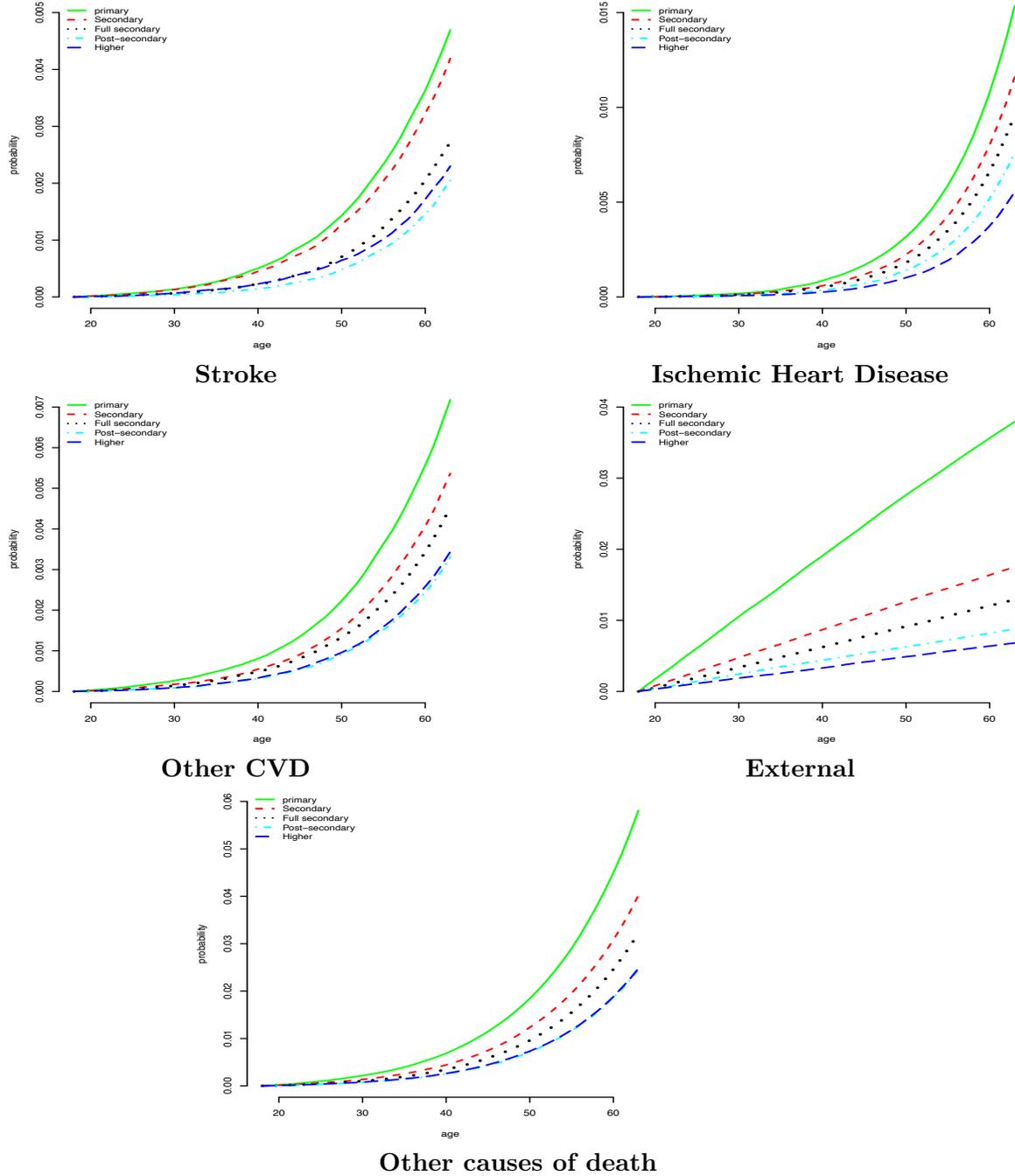


Figure D.9: Simulated total educational gain in cumulative incidence curves by cause of death (CVD) and education level

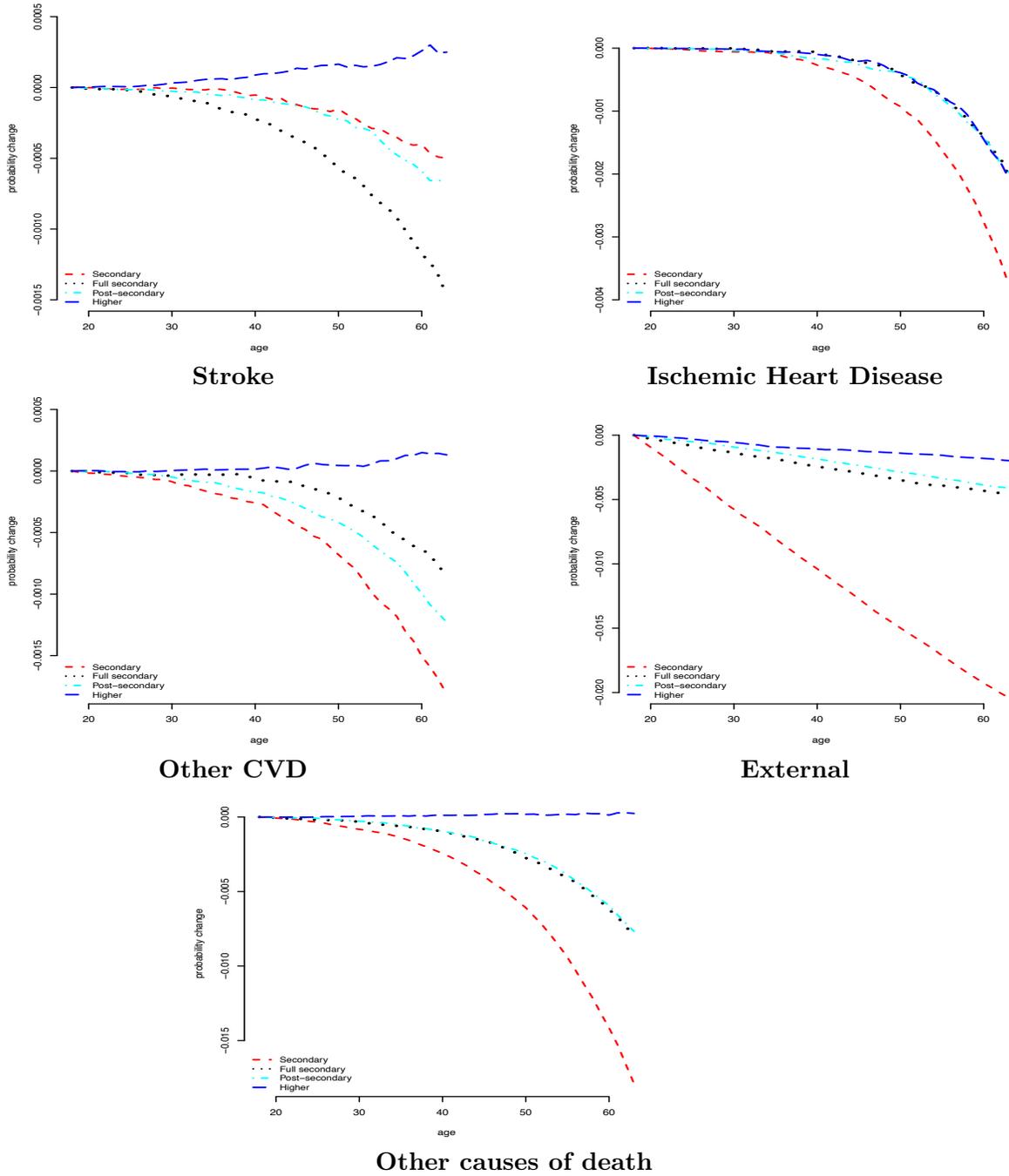


Figure D.10: Simulated Direct educational gain in cumulative incidence curves by cause of death (CVD) and education level

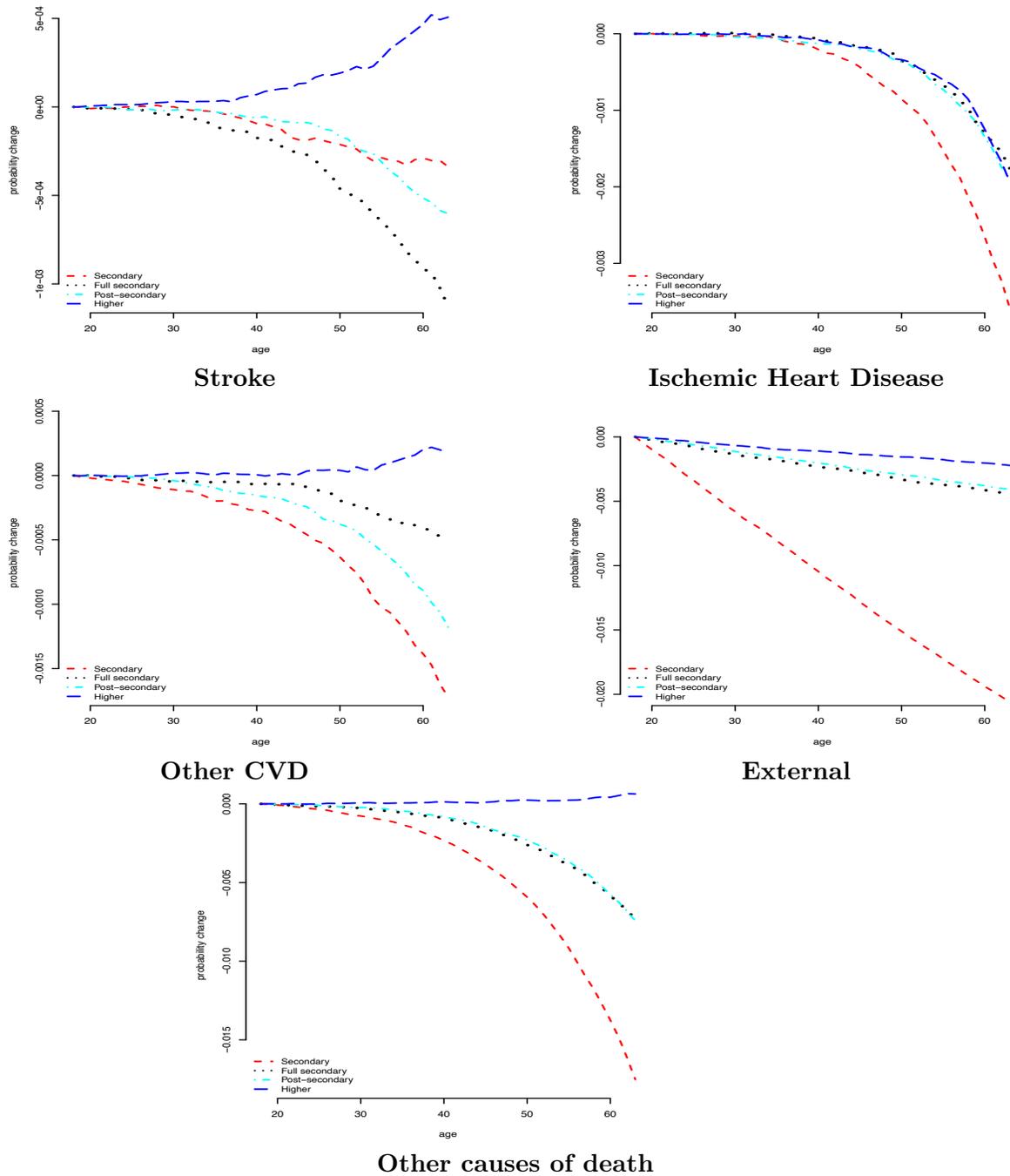


Figure D.11: Simulated Indirect effect of educational gain in Cumulative incidence curves by cause of death (CVD) and education level

