



K O N I N K L I J K E N E D E R L A N D S E
A K A D E M I E V A N W E T E N S C H A P P E N

Educational gain in cause-specific mortality: accounting for confounders

Bijwaard, G.E.; Myrskylä, M.; Tynelius, P.; Rasmussen, F.

2017

document version

Publisher's PDF, also known as Version of record

document license

CC BY-NC-ND

[Link to publication in KNAW Research Portal](#)

citation for published version (APA)

Bijwaard, G. E., Myrskylä, M., Tynelius, P., & Rasmussen, F. (2017). Educational gain in cause-specific mortality: accounting for confounders. (pp. 1-34). (MPIDR Working Paper; No. 2017 - 003). Max Planck Institute for Demographic Research .
http://www.demogr.mpg.de/En/projects_publications/publications_1904/mpidr_working_papers/educational_gain_in_cause_specic_mortality_accounting_for_confounders_5733.htm

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the KNAW public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the KNAW public portal.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

pure@knaw.nl



Max-Planck-Institut für demografische Forschung
Max Planck Institute for Demographic Research

Konrad-Zuse-Strasse 1 • D-18057 Rostock • Germany • Tel +49 (0) 3 81 20 81 - 0 • Fax +49 (0) 3 81 20 81 - 202 • www.demogr.mpg.de

MPIDR Working Paper WP 2017-003 | January 2017

Educational Gain in Cause-Specific Mortality: Accounting for Confounders

Govert E. Bijwaard | bijwaard@nidi.nl
Mikko Myrskylä | sekmyrskylä@demogr.mpg.de
Per Tynelius
Finn Rasmussen

© Copyright is held by the authors.

Working papers of the Max Planck Institute for Demographic Research receive only limited review. Views or opinions expressed in working papers are attributable to the authors and do not necessarily reflect those of the Institute.

Educational Gain in Cause-Specific Mortality: Accounting for Confounders

Govert E. Bijwaard^{*1}, Mikko Myrskylä², Per Tynelius³ and Finn Rasmussen³

¹Netherlands Interdisciplinary Demographic Institute

²Max Planck Institute for Demographic Research

³Department of Public Health Sciences, Karolinska Institute

January 2017

Abstract

For many causes of death a negative educational gradient has been found. This association may be partly explained by confounding factors that affect both education attainment and mortality. We correct the cause-specific educational gradient for observed individual background and unobserved family factors, using an innovative method based on months lost due to a specific cause of death re-weighted by the probability to attain a higher education level. We use men with brothers in the Swedish Military Conscription Registry (1951-1983), linked to administrative Swedish registers. These data, comprising 700,000 men, allow us to distinguish five education levels and many causes of death.

The empirical results reveal that improving education from primary to higher education would lead to 20 months longer survival between 18 and 63. The reduction in death due to external causes when improving education is attributable to most of these gains. Ignoring confounding would underestimate the educational gains, especially for the low educated. Implied by our results is that if 50,000 men from the 1951 cohort had had the 1983 education distribution they would have saved 22% of the person-years between ages 18 and 63 that were lost to death.

^{*}Netherlands Interdisciplinary Demographic Institute (NIDI-KNAW/University of Groningen), PO Box 11650, 2502 AR The Hague, the Netherlands, +31 70 3565224, bijwaard@nidi.nl

1 Introduction

Higher mortality rates among individuals with lower socioeconomic status are among the most compelling and well established facts in social science research (Hummer and Lariscy 2011; Mirowsky and Ross 2003). Among the different socioeconomic measures the educational gradient has been particularly robust (Cutler et al. 2011). This educational gradient is not only apparent in less-developed countries but also in the US (Masters et al. 2012) and other Western countries with advanced health care systems (Huisman et al. 2005). For example, in the year 2010, life expectancy at 30 was 53 years for Swedish men with a university degree and only 48 for Swedish men with less than secondary education. Educational differences have been reported for most causes of death, in particular for cardiovascular diseases and some types of cancer (Huisman et al. 2005). Many of these causes of death are theoretically avoidable through prevention or treatment. According to the ‘fundamental cause theory’ by Link and Phelan (1995) the educational gradient is larger for such causes of death, because personal resources and social context can be used to acquire health-related knowledge that prevents these causes, increases recovery (Link and Phelan 1995; Mackenbach et al. 2015). High educated individuals can handle complex treatments easier. According to this theory the educational gradient for non-preventable diseases, which are less under human control, is smaller.

It is commonly assumed that this educational gradient in mortality has a causal interpretation. However, this assumption has been challenged in the literature. The association may partly be explained by confounding factors, observed or unobserved variables, that affect both educational attainment and mortality (Grossman 2015). Ignoring that such common traits exists may render the association spurious and therefore change the policy implications. Still most of the studies investigating educational differences in cause-specific mortality do not account for the existence of confounding factors (Huisman et al. 2005; Kulhánová et al. 2014; Mackenbach et al. 2015).

In the literature three different approaches have been employed to correct for possible confounding factors. First, many studies rely on instrumental variable methods that use an exogenous variable that is correlated with education but not with the possible confounding factors. These studies usually exploit changes in compulsory schooling laws that increase schooling without (directly) influencing health. Results from recent analyses using such an instrumental variable method suggest that the causal effect of education on mortality may be limited (Mazumder 2008; Van Kippersluis et al. 2011; Meghir et al. 2013) or even absent (Albouy and Lequien 2009; Clark and Royer 2013). A second approach to account for confounding factors is to control for unobserved genetic and environmental factors by examining the effects of differences in schooling obtained by identical twins. Results from such studies indicate that part of the educational differences in mortality disappears when accounting for shared family background (Behrman et al. 2011; Lundborg 2013; Næss et al. 2012; Amin et al. 2015; Lundborg et al. 2016). A third approach to account for confounding factors is to include them directly into the model. A structural framework developed by Conti et al. (2010) explicitly models the interdependence between education, health, and the confounding factor(s). Results from such models for mortality (Bijwaard et al. 2015a,b) show that at least half of the health disparities across educational groups is due to the selection of healthier, more able individuals into higher education. In a recent paper we extended this framework to analyse the educational gains in cause-specific mortality (Bijwaard et al. 2016), revealing that the largest educational gains can be achieved by the low educated men in the reduction of external causes of death. However, these structural models 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 very computer intensive if you want to allow for many different causes of death.

The propensity score method (see Caliendo and Kopeinig (2008) for a survey) we employ in this paper also accounts for possible confounding factors, however, without making any structural assumptions on the relation between the confounding factors, education and (cause-specific) mortality. The advantage of the propensity score is that it enables us to summarize the many possible confounding

covariates as a single score. Propensity score weighting methods for hazard models, such as the mortality hazard, that account for censoring, truncation and dynamic selection issues have been introduced recently (Cole and Hernán 2004; Austin 2014). The critical assumption in propensity score weighting is that of no selection on unobserved factors. However, twin studies have emphasized that even after accounting for the socioeconomic background the association of education with mortality may be a reflection of other factors shared within a family. Apart from the 50% generic similarities siblings share environment during their childhood. To account for these unobserved family characteristics on the educational gradient in mortality we apply a family fixed effects model only comparing siblings (brothers).

The dominant way to analyze the educational impact on cause-specific mortality has been to estimate a Cox proportional hazard, Cox (1972), for each cause of death separately. This assumes that the competing causes of death are independent and leads to a situation where the removal of one cause will leave the risk of dying from other causes unchanged. Another issue is that in the presence of competing risks and when education influences each cause-specific hazard, the interpretation of the effect of education on the probability of dying from a specific cause is not obvious, see Thomas (1996). In such a competing risk model even the sign of the education effect is not directly clear, because both the total survival and the cause specific cumulative incidence functions not only depend on the cause specific hazard but also on the hazards of all other causes.

An alternative is to regress education on the cumulative incidence functions, the probability of dying from a specific cause before some age. Fine and Gray (1999) introduced a Cox-type model to regress directly on cumulative incidence functions, using subdistribution hazards. However, interpretation of the regression results is very difficult (Putter et al. 2007; Beyersmann and Scheike 2013). The contribution of a given cause in the educational gain cannot be derived from neither a Cox proportional hazard nor a Fine and Gray regression method, because they do not account for the importance of a specific cause of death. A direct way, that accounts for the importance, to measure the impact of education on cause-specific mortality is to estimate the impact of education on the months lost due to a specific cause. Andersen (2013) shows that, using pseudo-observations, this measure is easily adjusted to calculate the months lost before some pre-specified age (e.g. the maximum observed age of death in a sample). To account for observed confounding factors we estimate these month-lost models using inverse probability weighting (IPW) based upon the propensity score (Hirano et al. 2003). Weighting by the propensity score method creates a synthetic sample in which the educational attainment is independent of the included covariates. Based on this synthetic sample we estimate linear month-lost models with family fixed effects. The family fixed effect account for unobserved family confounding factors.

Data from the Swedish Military Conscription Register for men born 1951-1983, linked to Swedish administrative registers, offers the opportunity to investigate these questions. These data include recording of demographic and socioeconomic characteristics such as education, parental (both fathers and mothers) socioeconomic status, parental education, area of residence along with anthropometric measures, an intelligence test and a psychiatric assessment. Educational level was classified in five categories: primary education; secondary education (2 years); full secondary education (3 years); post-secondary education and higher education. We have sufficient data to distinguish 22 different causes of death.

A sample of only those men who were born in 1951-1960 are used in Bijwaard et al. (2016). In that paper we also investigate the educational gains in cause-specific mortality, based on a structural model instead of an IPW-estimation method. In this paper we use all the available birth cohorts, but limit the analyses to men with brothers, and have an almost twice as large sample of 700,043 men with known conscription date (and at least one brother). The large sample allows us to distinguish 22 instead of four (major) causes of death and five instead of four education levels.

The empirical results show that improving education one level would lead to three to ten (primary education) additional months alive between 18 and 63 (or 20 months in total when improving education

from primary to higher education). Most of these gains is attributable to the reduction in death due to external causes. Men with primary education would gain the most from the reduction in suicide (1.3 months) and other external death (2.4 months). The educational gain for CVDs and cancers is rather small. Ignoring confounding would underestimate the educational gains for the low educated (primary education).

Implied by our results is that if 50,000 men from the 1951 cohort had had the 1983 education distribution they would have saved 423 person-years between age 18 and 63, which is 22% of the person-years between ages 18 and 63 that were lost to death. The main source of these saved person-years is a reduction of 248 person-years in death due to external causes (28% reduction). Reduction in mortality due to cancer, saving 28 person-years (8% reduction) and, due to cardiovascular diseases, saving 53 person-years (18% reduction), would only play a minor role. Finally, reduction in mortality from other natural causes would save 94 person-years (28% reduction).

The paper is structured as follows. In Section 2 we discuss relevant previous research and conceptual framework. In section 3 we explain the use of pseudo-observations to allow for interaction among the causes of death and the use of inverse propensity score weighting to correct for selection into education. Section 4 presents data on education, confounding variables and mortality outcomes. Section 5 presents the results and Section 6 discusses them.

2 Previous Research and Conceptual Framework

A conceptual challenge exists in defining socioeconomic status. Often used measurements include educational attainment, income or occupational class. It has been demonstrated that these factors cannot be used interchangeably (Geyer et al. 2006). Educational status is the most commonly used indicator of socioeconomic status in studies of health and mortality (Hummer et al. 1998; Preston and Taubman 2011; Hummer and Lariscy 2011), as education precedes both income and occupation (Ross and Wu 1995; Mirowsky and Ross 2003). Another advantage of using educational status is that education is usually completely early in life and remains constant over the life course and does not respond to health fluctuations (Smith 2004).

Most studies focussing on the educational gradient in health and mortality measure educational attainment with a single indicator of years of completed schooling assuming that each additional year of education confers a monotonic increase in health (Elo and Preston 1996; Lynch 2003). We assume that the educational gradient is a step function in the degree earned (Backlund et al. 1999; Montez et al. 2012), distinguishing five ordered education levels *i*) Primary school (less than 10 years of primary school); *ii*) Secondary education (at most 2 years); *iii*) Full secondary education (2-3 years); *iv*) Post-secondary education (less than 3 years) and *v*) Higher education (University and PhD).

Cause specific mortality

For most causes of death large mortality differences between individuals in low- and high socioeconomic position have been observed in both North America and Europe. The persistence of these inequalities in health is one of the most frustrating disappointments of public health in these highly developed countries. A couple of theories explaining these health inequalities exist (Mackenbach 2012), we review the most important ones. According to the ‘fundamental cause theory’ (Link and Phelan 1995) the educational gradient should be the largest for the causes of death that are more preventable and/or curable. This theory stipulates that these health differences are caused by socioeconomic differences in access to resources which can be used to avoid these diseases and slow down disease development once the diseases occurs regardless of the current exposure to risk factors. The ‘life course perspective’ emphasizes the importance of unfavorable early life circumstances in explaining the pathways to both health and social disadvantage in adulthood. The widespread positive education-mortality associations

may not necessarily reflect beneficial effects of education on mortality, because other individual factors may influence both education and mortality.

Evidence suggests differential impact of education on various diseases accumulating in different educational cause-specific mortality gradients, Galobardes et al. (2004). As the socioeconomic association seems the largest for cardiovascular diseases and some types of cancer most studies have focused on socioeconomic differences in mortality on these type of diseases. The incidence of cardiovascular disease is indeed higher for individuals with low socioeconomic status (Mackenbach et al. 2008). The educational gradient for cardiovascular diseases (CVD) appears to be stronger than for total mortality (Kulháňová et al. 2014). The most probable reason is that low education has been associated with cardiovascular risk factors, such as smoking, hypertension and overweight. For cancers the education-gradient varies by cancer type (Galobardes et al. 2004; Kulháňová et al. 2014). Higher mortality for the low educated from lung cancers are clearly related to the higher smoking prevalence of these individuals. The relationship between education level and the mortality rate for other cancers is more complex. Lifestyle differences, such as physical inactivity, might be one reason for this. However, a recent study found little evidence that education affects cancer mortality (Leuven et al. 2016), except for lung and prostate cancer for men.

Only a few studies have investigated the relation between education and death from external causes, including suicides and traffic accidents (Borrell et al. 2005; Lorant et al. 2005). Such mortality is a major cause of early (young adult) death. Socioeconomic inequalities in suicide are observed in many countries for males. Educational differences in mental illness, which is more prevalent among the low educated, may explain this educational gradient. The educational differences in traffic accidents can be explained by differences in exposure, such as different use of protective devices, and differences in susceptibility.

A major limitation of most of the studies analysing educational differences in cause specific mortality is that they only report associations. Only a few recent papers have investigated the causal effect of education on specific causes of death.

Causal inference

We seek to find the causal impact of education level on cause-specific mortality risk. 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, usually increases in the minimum age or the legally permitted grade to leave school, as instrumental variables for schooling attainment to control for possible confounders, factors that affect both the education attained and mortality. The estimates based on these studies point towards a small effect (Lleras-Muney 2005; Van Kippersluis et al. 2011; Meghir et al. 2013), or even entirely absent (Arendt 2005; Albouy and Lequien 2009; Clark and Royer 2013) causal effect of education on health outcomes. 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 they may accompany these reforms.

A second approach to account for confounding factors is to control for unobserved genetic and environmental factors by examining the variation in education among siblings, often identical (monozygotic) twins. These studies obtain estimates of the impacts of the differences in schooling within a pair of identical twins on their health differences at various schooling levels. Results from such studies indicate that part of the educational differences in cause-specific mortality disappears when accounting for shared family background (Behrman et al. 2011; Lundborg 2013; Næss et al. 2012; Amin et al. 2015). These sibling studies also suffer from limited data and issues of generalization, as they only use data on siblings and those families might not represent the general population.

A third approach to account for confounding factors is to include them directly into the model. A structural framework developed by Conti et al. (2010) explicitly models the interdependence between education, health, and the confounding factor(s). Results from such models for mortality (Bijwaard et al. 2015a,b) show that at least half of the health disparities across educational groups is due to the selection of healthier, more able individuals into higher education. In a recent paper we extended this framework to analyse the educational gains in cause-specific mortality (Bijwaard et al. 2016), revealing that the largest educational gains can be achieved by the low educated men in the reduction of external causes of death. However, these models 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 you want to include many different causes of death.

The propensity score method we employ in this paper also accounts for possible confounding factors, however, without making any structural assumptions on the relation between the confounding factors and (cause-specific) mortality. It is based on the assumption that all variables that affect both mortality 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). We also relax this assumption by additionally comparing siblings (brothers) and use a propensity score method with fixed effects that allows for unobserved family characteristics on the educational gradient in mortality.

3 Methodology

The standard approach to analyse cause-specific mortality is to formulate a competing risks model with an independent Cox proportional hazards model for each cause-specific outcome. The cause-specific hazard gives the mortality rate due to a particular cause conditional on not having died from any other cause previously. The interpretation of the coefficients, or hazard ratios, in such a competing risks model requires caution (Thomas 1996). A particular covariate can appear in several competing hazards. In such a case the coefficients (or hazard ratios) related to this covariate convey little information about the effect of the covariate on the probability to die from a specific cause, because the probability to die from a specific cause not only depends on the hazard to die from that cause but also on the hazards to die from all the other causes. Thus, despite that in a Cox analysis for competing causes of death it is assumed that the causes are independent, many measures of the importance of a specific cause depend on the hazards of the other causes. A final objective against using Cox competing risk models is that, for a given cause of death, they only give the magnitude of the educational disparity and not the importance of the specific cause.

Another approach to model competing risks is to estimate a regression model for the cumulative incidence function, which is the probability of dying from a particular cause before some age t . The Fine and Gray (1999) regression model for the cumulative incidence assumes that the cumulative incidence for a particular cause is given by a complementary log-linear model. This model has some resemblance with the Cox proportional hazards model. However, the interpretation of the coefficients is rather difficult, because the Fine–Gray model is based on the ‘subdistribution hazard’, which is not the cause-specific hazard. Implicit in the subdistribution hazard is that it is assumed that any individual who dies from a competing cause remains in the risk set for a specific cause. This implies that the risk sets associated with the subdistribution hazard will be unknown for competing events after they have occurred.¹

¹For a detailed discussion on these issues see Putter et al. (2007) or Beyersmann and Scheike (2013).

3.1 Months lost due to a specific cause of death

A quantity with a more natural interpretation that avoids the issues of interdependence in a competing risks proportional hazards model is the number of *months lost* due to a specific cause of death, see Andersen (2013) and Andersen and Canudas-Romo (2013). The months lost can be defined over the whole age distribution, e.g. the number of months lost before age 63 (as we will use). A nice feature of the months lost quantity is that the sum over all alternative causes of death is equal to the total amount of months lost. It can be calculated using non-parametric methods and based upon estimated hazard coefficients, the implied total survival and the cumulative incidence functions. Non-parametric estimation of the months-lost is straightforward because the survival (Kaplan-Meier) and cumulative incidence functions are straightforward to estimate, see Appendix A.

The advantage of months-lost is that they can be modelled using standard (uncensored) linear models. However, a common issue with mortality data is that not all individuals are followed till they die, but only till the end of the observation period. This implies that the age at death is (heavily) censored, in our case at age 63. This complicates regression analyses. With constructed pseudo-observations it is possible to carry out regression analyses for the months-lost for each specific cause of death when some of the observations are censored, see Andersen (2013).²

3.2 Accounting for confounding

We seek to find the causal impact of education on cause-specific mortality. However, (cause-specific) mortality may be influenced by factors that also determine the educational attainment. This may render educational attainment endogenous to mortality later in life.

We follow a propensity score method to account for this endogeneity of education. Propensity score methods are increasingly used to estimate causal effects in observational studies, e.g. see Caliendo and Kopeinig (2008) for a survey. These methods aim to adjust for confounding factors between the treatment groups, in our case different education levels. The advantage of the propensity score is that it enables us to summarize the many possible confounding covariates as a single score (Rosenbaum and Rubin 1983). Propensity score methods include matching, stratification on the propensity score and inverse probability weighting (IPW) using the propensity score. The methods we apply are based on IPW-methods (Hirano et al. 2003).

We define the educational gain for each cause of death in terms of the change in months-lost. The advantage of using a months-lost model is that the total gain of moving up the whole educational ladder (from primary education to higher education) is the sum of each one-level educational gains. Common assumptions in the literature using propensity score methods to identify these ‘treatment effects’ are the *Unconfoundedness* and the *common support* assumptions. The unconfoundedness assumption (Rubin 1974, Rosenbaum and Rubin 1983) asserts that, conditional on observed individual characteristics education attainment is independent of the *potential* outcomes (months lost). This implies that (conditional on observed characteristics) the difference in the *potential* outcome if the individual had had low education and the *potential* outcome if the individual had had high education is only caused by education. This assumption requires that all variables that affect both mortality and education attainment are observed. Note that this does not imply that we assume all relevant covariates are observed. Any missing factor is allowed to influence either mortality or education attainment, not both. Although this is not testable and clearly a strong assumption, it may be a reasonable approximation. Any alternative, that does not rely on unconfoundedness while allowing for consistent estimation of the educational gain, will have to make alternative untestable assumptions. The overlap, or common support, assumption requires that the propensity score, the conditional probability to attain higher education given observed individual characteristics, is bounded away from zero and one. This assumption is in principle testable. If there are values of the included covariates

²We explain this method in more detail in Appendix A.1.

for which the probability of observing a higher education level is zero or one, we cannot compare these individuals between a high and a low education level. In that case we have to limit comparisons to sets of values where there is sufficient control in the propensity score among both low- and high educated. By comparing only adjacent education levels we remove the overlap problems.

If unconfoundedness holds, all biases due to observable covariates can be removed by conditioning on the propensity score (Imbens 2004). The average educational gains can be estimated by weighting on the propensity score. Inverse probability weighting based on the propensity score 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.

We implement an IPW-version of the estimation procedure for months-lost based on pseudo-observations. Xiang and Murray (2012) developed a related method for estimation of the restricted mean. The estimation of the IPW-months lost is straightforward and based on an extension of the Stata procedure `stplost` of Overgaard et al. (2015). After calculating the weights, based on the estimated propensity scores, we construct a pseudo population. Then, for this pseudo population we estimate the total survival-function, using a Kaplan-Meier estimate, and the cumulative incidence function, using the Aalen-Johansen method, and use these two quantities to estimate the base months-lost for each cause of death. Next we regress these IPW-pseudo observations on the education indicator to obtain our estimate of the impact of education on months-lost.

One of the weaknesses of previous analyses is that estimated educational effects may be biased by unobserved genetic or social factors that influence both the education level and mortality. To eliminate this bias we also include family fixed effects in the month-lost models.

4 Data

The data was retrieved from several Swedish population-wide registers which were linked using a unique individual identification number assigned to all Swedish citizens. The Swedish Military Conscription Register (SMCR) includes demographic information (birthdate, area of residence) intelligence tests, a psychological assessment and anthropometric and health measures (height, weight, blood pressure and a muscular strength test). This SMCR was linked to the National Population and Housing Censuses (1960-1990) with information on the parental socioeconomic status and education. The conscripts own education was obtained from the Longitudinal Integration Database for Health Insurance and Labour Market Studies (LISA) during 1990-2012, and cause-specific mortality as the underlying cause of death from the Cause of Death Register with follow-up till 2012.

The study population consists of men born between 1950 and 1983 identified in the Multi-Generation Register, and who were enlisted between 1969 and 2001 in the year they turned 18-20. In Sweden military service was mandatory only for men. We selected only those men for whom at least one parent is known. We also removed men without a known conscription date. Finally, because we want to account for shared family influences, we only consider men with (observed) brothers.³ We aggregated the observed education into five classes: (*i*) primary education (less than 10 years of education); (*ii*) 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). We identified 700,043 men with at least one brother in 369,526 families of which 13% belongs to the lowest education group; 31% has finished secondary education (2 years); 20% has finished full secondary education (3 years); 15% has post-secondary education and 21% has attained higher education.

Selected demographic and socioeconomic characteristics at the time of military examinations by

³We did not identify substantial differences in observed individual characteristics between men with and without brothers.

education level are given in Table 1⁴. We see a clear relation between parental socioeconomic status (SES), parental education and education attained by the conscript. The higher the social status and the education of the parent the higher the education of the conscript. Intelligence and psychological fitness are also clearly related with the attained education.

⁴In Table B.2 in the appendix we give the descriptives for the men without a (known) brother.

Table 1: Descriptive statistics

	Primary	Secondary education (2 years)	(3 years)	Post-secondary (< 3 years)	Higher
<i>SES mother at birth</i>					
non-manual (high)	1%	1%	2%	4%	8%
non-manual (intermediate)	3%	4%	9%	12%	19%
non-manual (low)	13%	18%	20%	28%	28%
Farmers	13%	9%	5%	6%	4%
Skilled workers	39%	38%	16%	20%	11%
Unskilled workers	15%	16%	18%	13%	10%
Not classified	16%	13%	30%	16%	20%
Unknown	1%	1%	1%	1%	1%
<i>Education mother</i>					
Primary (< 9 yrs)	54%	45%	22%	23%	13%
Primary (9–10 yrs)	12%	11%	16%	11%	9%
Secondary education (2 yrs)	25%	32%	38%	33%	28%
Secondary education (3 yrs)	2%	3%	6%	7%	8%
Post-secondary	3%	5%	10%	13%	18%
Higher	2%	3%	7%	12%	23%
Unknown	2%	1%	1%	1%	1%
<i>SES father at birth</i>					
non-manual (high)	2%	2%	5%	9%	18%
non-manual (intermediate)	7%	11%	17%	25%	29%
non-manual (low)	7%	8%	11%	11%	11%
Farmers	9%	7%	7%	5%	4%
Skilled workers	28%	30%	27%	21%	15%
Unskilled workers	40%	37%	28%	22%	15%
Not classified	6%	5%	5%	6%	7%
Unknown	1%	1%	1%	1%	1%
<i>Education father</i>					
Primary (< 9 yrs)	58%	51%	29%	27%	16%
Primary (9–10 yrs)	7%	6%	12%	6%	5%
Secondary education (2 yrs)	17%	21%	26%	21%	17%
Secondary education (3 yrs)	7%	11%	16%	20%	18%
Post-secondary	2%	4%	7%	10%	11%
Higher	2%	3%	8%	15%	31%
Unknown	6%	3%	2%	2%	2%
mother < 20 at birth	10%	9%	5%	5%	3%
father > 40 at birth	8%	7%	5%	6%	5%
birth order	3.1	2.9	2.8	2.6	2.5
global IQ ^a	3.7	4.4	5.0	6.1	6.6
Psychological assessment ^a	4.1	4.7	5.1	5.6	5.8
# of individuals	91,588	219,372	138,653	101,883	148,547

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

5 Results

Before we turn to the estimation results from the month-lost models we discuss the results from standard Cox proportional hazards models, the commonly applied model in the literature to analyse cause-specific competing risks (Elo et al. 2014; Kulhánová et al. 2014). However, in using the Cox models we deviate from the standard literature and instead of including four educational dummies, one for each of the four education levels above primary education, we estimate four separate models using the data of individuals in two adjacent education levels (for each cause of death) with only one dummy variable for the higher educational level. This relaxes the assumption of common age-dependence by educational level, implicit in the Cox model, but still ignores that both the attained education and the causes-specific mortality may depend on the same individual characteristics.

Table 2 reports the estimated hazard ratios of education (moving up one level) for each cause-specific outcome.⁵ We find that for many causes of death, except for many cancers, an improvement in education reduces the risk of dying from that cause. In the comparison of the lowest two education levels only the mortality risk of some cancers, stroke and infectious diseases are not significantly lower for men with secondary education and, the largest educational impact is found for leukemia for which the mortality rate is more than halved. When comparing men with 2 years and 3 years of secondary education none of the cancer mortality rates and half of the mortality rates for other diseases differ significantly by education level and, the largest educational impact is found for infectious diseases. When comparing men with 3 years of secondary education and post-secondary education digestive cancer is the only cancer with a significant educational difference in mortality. The largest impact of post-secondary education on the mortality rate is found for psychiatric diseases. Finally, for the comparison between the highest two education levels, post-secondary- and higher education, we find a significant educational impact for two of the cancer mortality rates, for a few other diseases (stroke, other CVD and, other natural causes) and, for mortality due to external causes. The largest impact of higher education on cause-specific mortality is found for other natural causes.

Next, we estimate the educational impact on the months-lost due a specific cause of death. The months lost model accounts for possible interdependence of the mortality rates and it also warrants easy interpretation. We focus on the months lost lived from 18, the age at military examination, till 63, the maximum age observed. The columns denoted by (1) in Table 3 provide the estimated results. The ‘primary’ column reports the estimated months lost for the men with only primary education. The next columns report the educational gain in months lost for each additional education level. An advantage of the month-lost approach is that we can sum the months-lost (and educational gains) over the causes of death (and over the education levels). This allows us to directly derive the educational gains on the major causes of death (grouped), cancers, CVDs, other diseases and, external causes and the educational gain when improving education over the whole education ladder from primary education to higher education.

According to the results men with only primary education loose in total 23 months of their life between age 18 and 63. Improving education to secondary education reduces the months lost with 8.5 months. Getting one additional year of secondary education reduces the months lost with an additional 2.9 months, obtaining post-secondary education reduces the months lost 3.5 further and, finally, obtaining higher education further reduces the months lost with 2.5 additional months. Although (almost) the same causes of death as in the Cox proportional hazard analyses show a significant educational gain, the importance of specific causes changes. Only from the month-lost analyses we can conclude that the three external causes of death are far more relevant (for mortality between 18 and 63) in explaining the educational gain (and in explaining the total mortality) than other causes of death. For example, the amount of months lost due to suicide decreases with 1.4 months when improving education to secondary education (2 years) and only with 0.6 month for all cancers together

⁵To check whether men with brothers are a selective group of all conscripted men we also estimated Cox models and months lost models for the full sample of men, with or without brothers, see Table B.1 and Table B.3 in appendix B.

Table 2: Cox hazard estimates education impact cause-specific mortality (HR), for men with brothers

	Secondary education (2 years)	Post-secondary (3 years)	Higher (< 3 years)	Higher
<i>Cancers</i>				
bowel cancer	0.951	0.824	0.898	1.171
digestive cancer	0.859	0.963	0.661**	0.734
lung cancer	0.733 ⁺	0.932	0.724	0.551 ⁺
leukemia	0.444**	1.025	1.262	0.497 ⁺
lymphoma	1.113	0.841	1.226	0.829
other cancer	0.903	0.918	0.925	0.867
<i>CVD</i>				
Ischemic heart disease	0.656**	0.794**	0.631**	0.836
stroke	0.906	0.647**	0.799	0.569 ⁺
other CVD	0.774**	0.908	0.702**	0.460**
<i>Other diseases</i>				
Infectious	0.795	0.389**	0.855	1.058
endocrine	0.671	0.800	1.455	0.607
diabetes	0.618**	0.628 ⁺	0.862	0.598
psychiatric	0.491**	0.555**	0.261**	0.615
nervous system	0.514**	0.995	0.759	0.652
respiratory disease	0.570**	0.397**	0.439 ⁺	2.024
digestive disease	0.490**	0.643	0.886	0.593
liver	0.702**	0.520**	0.427**	0.767
abnormalities	0.569**	0.900	0.685	0.988
other natural causes	0.463**	1.295	0.583	0.227**
<i>External causes</i>				
traffic accidents	0.776**	0.698**	0.634**	0.431**
suicide	0.700**	0.723**	0.763**	0.592**
other external causes	0.572**	0.652**	0.670**	0.444**

Separate models are estimated for each two adjacent education levels. ⁺ $p < 0.05$, ** $p < 0.01$.

or 1.3 month for all cardiovascular diseases together. The relative importance of external causes in the total months-lost gain is for all education levels above 50%. Although the highest education group only gains 1.8 months due to external causes they comprise 71% of the educational gain. Other important causes of death are ischemic heart diseases, psychiatric diseases (except for the highest education level) and, other CVDs.

However, it is very likely that education attainment and cause-specific mortality both depend on individual characteristics. Ignoring this interdependence, confounding, may bias the estimated educational gains.

5.1 Adjusting for confounding and family fixed effects

To adjust for such confounding we re-estimate the month-lost models for each specific cause of death using a re-weighted pseudo-population based on inverse propensity score weighting, see Appendix A. To calculate the propensity score we could, in principle, estimate an ordered probit or ordered logit propensity score for our five ordered education levels, see Imai and van Dyk (2004) and Feng et al. (2012). However, the men in the lowest and highest education group differ too much in their observed background characteristics, which causes severe overlap problems. We therefore estimate separate

propensity scores of attaining a higher education level through pairwise comparisons (Lechner 2002) of adjacent education levels. We include variables that influence both the probability to obtain a higher education level and the probability to die. The control variables in the propensity score include parental (mothers and fathers separately) social economic status (SES) at birth, parental education, whether the mother was young or the father was old at birth, IQ level, psychological assessment, birth order, county dummies and, year of birth dummies.

The table with coefficients of the four estimated logit models can be found in the appendix, see Table B.4. Here we only discuss the main factors that influence the education attained. Not surprisingly, the intelligence of the individual plays an important role in explaining educational attainment, with a high IQ leading to higher education. It might be argued that intelligence is influenced by the education attained (Ceci 1991). However, Bijwaard and Jones (2016) show that treating intelligence as a mediator in the path from education to mortality gives very similar results as treating intelligence as one of the selection variables determining the educational attainment. A better psychological assessment is associated with higher education. Parental socioeconomic status has, in general, the expected influence on education attainment, with higher parental socioeconomic status leading to higher education of their offspring. Parental education has a similar influence on the attained education. Sons born when their mother was young have a lower education, and sons from older fathers have a higher education. We tested whether the propensity score is able to balance the distribution of all included variables in each of the two adjacent education groups using standardized bias calculations and we also checked for common overlap issues.⁶

Based on the logit estimates to attain a higher education level we calculate the propensity scores and weight the men with the higher education by the inverse of the propensity score and those with the lower education by the inverse of one minus the propensity score (within each pair of education levels in the four education groups). However, only accounting for observed confounders is not sufficient as unobserved genetic or social factors that influence both education level and mortality may still exist. To control for this bias, we also account for family fixed effects in the weighted months lost models.

The results from these models are presented in the columns denoted by (2) in Table 3. Accounting for both observed confounders, through the IPW method, and unobserved shared family factors, through fixed effects, increases the educational gain for all education groups, especially for men with only primary education. For all education groups the gain due to CVDs increases after accounting for confounding factors. Still, external causes are attributable to the largest gain in months lost.

To illustrate the size of the estimated educational gains we calculate how many men-years the 50,000 men born in 1951 would gain if they had the 1983 education distribution. Based on the IPW with family fixed effects results the 1951-cohort would gain 423 person-years if they had had the 1983 education distribution, which is a 22% reduction of the person-years between ages 18 and 63 that were lost to death. Most of this gain, 248 person-years (28% reduction), is attributable to reduced mortality due to external causes (traffic accidents 57 person-years, suicide 88 person-years and other external causes 103 person-years). The contribution of the reduction in cancer-mortality (28 person-years, 8% reduction) and CVD-mortality (53 person-years, 18% reduction, of which 27 person-years due to IHD) is rather small. The reduction of death to other diseases leads to 94 person-years (28% reduction) saved (30 person-years due to psychiatric diseases, 40% reduction).

⁶See appendix B.

Table 3: Months lost due to a specific cause 18-63 by education level, for men with brothers. (1) unadjusted model; (2) IPW model with family fixed effects

	Primary		Educational gain				Post-secondary (< 3 years)		Higher			
	(1) ^a	(2) ^a	(2 years)		(3 years)		(< 3 years)		Higher			
			(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)		
<i>Cancers</i>												
bowel cancer	3.28**	3.49**	0.61**	0.98 ⁺	0.20	0.45	0.28 ⁺	0.57	0.35**	0.44		
digestive cancer	0.34**	0.35**	0.02	0.04	0.05	0.07	0.02	0.02	-0.03	-0.02		
lung cancer	0.72**	0.70**	0.14	0.16	0.03	0.12	0.18 ⁺	0.22	0.05 ⁺	0.01		
leukemia	0.45**	0.42**	0.16 ⁺	0.14	0.01	-0.16	0.08	0.15	0.05 ⁺	0.11		
lymphoma	0.28**	0.26**	0.18**	0.13	-0.00	-0.00	-0.03	-0.03	0.07 ⁺	0.04		
other cancer	0.33**	0.34**	-0.05	-0.04	0.03	0.02	-0.04	-0.08	0.08	-0.03		
<i>CVD</i>	1.16**	1.43**	0.16	0.55 ⁺	0.07	0.40	0.07	0.29	0.12	0.32		
ischemic heart disease	3.67**	4.06**	1.33**	2.07**	0.38**	0.74	0.63**	0.90 ⁺	0.31**	0.72		
stroke	2.00**	2.22**	0.88**	1.30**	0.18 ⁺	0.37	0.33**	0.47 ⁺	-0.01	0.29		
other CVD	0.44**	0.51**	0.07	0.18	0.12**	0.26	0.07	0.12	0.05	0.16		
<i>Other diseases</i>	1.23**	1.34**	0.38**	0.59**	0.06	0.11	0.21	0.31	0.26**	0.27		
Infectious	4.98**	5.07**	2.53**	2.80**	0.67**	0.47	0.75**	0.59	0.11	0.25		
endocrine	0.30**	0.29**	0.11 ⁺	0.11	0.10**	-0.02	0.03	0.06	0.00	0.10		
diabetes	0.15**	0.14**	0.06 ⁺	0.04	0.01	0.04	-0.03	0.09 ⁺	0.04	0.08		
psychiatric	0.38**	0.33**	0.18**	0.11	0.07 ⁺	-0.01	0.03	0.01	0.03	0.07		
nervous system	1.27**	1.45**	0.74**	1.01**	0.23**	0.32	0.23**	0.09	-0.01	0.01		
respiratory disease	0.61**	0.66**	0.33**	0.41**	-0.02	0.05	0.08	0.08	0.06 ⁺	0.11		
digestive disease	0.60**	0.60**	0.34**	0.35 ⁺	0.13**	0.21	0.09 ⁺	0.03	-0.07	-0.11		
liver	0.21**	0.20**	0.13**	0.12	0.02	0.04	0.02	0.04	0.01	0.01		
abnormalities	0.67**	0.51**	0.24**	0.07	0.17**	-0.04	0.14**	-0.01	-0.02	-0.01		
other natural causes	0.59**	0.68**	0.29**	0.44**	-0.01	-0.12	0.10	0.15	0.02	-0.04		
<i>External causes</i>	0.20**	0.22**	0.11 ⁺	0.14	-0.03	0.00	0.06 ⁺	0.06	0.04 ⁺	0.03		
traffic accidents	10.70**	11.21**	4.10**	4.15**	1.71**	1.87 ⁺	1.86**	1.61	1.81**	1.60**		
suicide	2.24**	2.37**	0.61**	0.49	0.36**	0.43	0.59**	0.12	0.54**	0.34 ⁺		
other external causes	4.02**	4.04**	1.41**	1.27**	0.62**	0.59	0.60**	0.67	0.65**	0.71**		
total	4.44**	4.80**	2.04**	2.39**	0.71**	0.85 ⁺	0.65**	0.83 ⁺	0.61**	0.55 ⁺		
total	22.63	23.84	8.53**	9.99**	2.88**	3.52**	3.53**	3.67**	2.54**	3.00**		

^a (1) Unadjusted months lost model; (2) IPW with family fixed effects months lost model.

⁺ $p < 0.05$, ** $p < 0.01$.

6 Conclusion and discussion

Many studies have found a large positive association between education and longevity. Previous evidence suggests that the impact of education on various diseases differs. This accumulates in educational gradients that diverge by cause of death. Large educational differences in cardiovascular diseases have been reported, which is in line with the ‘fundamental cause theory’ that preventable diseases should exhibit larger educational gains. A major limitation of most of these studies is that they only report associations and ignore that part of the association between education and cause-specific mortality may be driven by confounding factors, factors that influence both educational attainment and cause-specific mortality. After accounting for both pre-educational individual and shared family factors we find, however, rather small educational differences in cardiovascular mortality and death due to other preventable diseases. We find the largest educational gains in the reduction of death due to external causes.

In contrast to common literature we analyse the months lost due to a specific cause of death instead of the cause specific (proportional) hazards. Standard Cox models assume independence of the competing risks and interpretation of the estimated education coefficients can be difficult. Another issue is that from Cox models only it is not possible to judge the importance of the cause in the total mortality. Using months lost due to a specific cause of death avoids these problems. Another advantage is that the month-lost results are, in contrast to hazard ratios, additive, both by education and by causes of death.

The inverse probability weighting (based upon the propensity score) method we employ accounts for observed confounding factors and unobserved shared family factors without making any assumptions on the relation between these factors and mortality. Thus, an advantage of our method is that we do not need to make any functional form assumptions of the impact of included control variables on mortality. Using a propensity score methods account for observed factors that influence both education and mortality. By focusing on men with brothers we were also able to adjust for unobserved family factors that remain the same across different births within a family, i.e. so called family fixed effects.

Based on standard Cox proportional hazard models we find, in line with the literature, many significant educational differences in cause-specific mortality. Using months-lost analyses and accounting for confounding factors reduces the number of causes with a significant educational gradient. In contrast to previous findings we hardly obtained any significant educational gains for death due to cancers or cardiovascular diseases. Importantly, we found that death due to external causes played a much larger role in explaining the educational differences. We found the largest educational gains for the lowest education group.

Based on our results we derive that if assuming that the 50,000 men of the 1951-cohort had had the 1983-cohort education distribution this cohort would have gained 423 person-years, which is a reduction in person-years lost of 22%. Most of this gain (248 person-years) is attributable to reduced mortality due to external causes and only a small part of the gain was due to a reduction in cancer-mortality (28 person-years) or cardiovascular diseases-mortality (53 person-years). The reduction of deaths to other diseases leads to 94 person-years saved.

Ignoring confounding factors would underestimate the educational gains for the low educated, especially with for men with only primary education. This implies that for the lower end of the education distribution men with secondary education are negatively selected (compared to men with only primary education), have higher mortality, This selection is mainly driven by differences in unobserved shared family unobserved factors, see Table B.6. This negative selection is most prominent for death due to CVDs. The negative selection hinges to lower education in families with lower mortality. Observed family factors that raise education, such as high parental SES and education, seem to play a minor role in explaining the gains in education.

Our study has four distinct strengths compared to previous research. First, a clear advantage of

the study is the very large sample size (700,000), which allows the estimation of the educational gain for five education levels and twenty-two causes of death. Second, the data are population based and not prone to self-selection issues because military conscription was mandatory in Sweden during the study period. Third, the statistical method we use, inverse propensity score weighting with family fixed effects, accounts for confounding effects on cause-specific mortality. This enables us to draw more accurate causal conclusions, without suffering generalization issues inherent to using compulsory schooling reforms to account for confounding. Fourth, contrary to the standard literature on causes of death (competing risks) analysis we define the educational gains of causes of death in terms of months lost due to each specific cause of death instead of the hazard ratio. This quantity has a more natural interpretation, provides a direct measure of importance of the specific cause, and avoids the issues of dependence in competing risks proportional hazard models. The months lost can be defined over a segment of the age distribution. The months lost quantity is an additive measure. The sum over all alternative causes of death within one education level is equal to the total amount of months lost (and the educational gain) for that education level and the sum of educational gains over all education levels within a cause of death is equal to the total educational impact of that cause of death.

A limitation of our data, based on military entrance examination, is that we do not have information on women. Another limitation is that we only observe mortality before the age of 63 (or earlier for cohorts from the 70s and 80s). In the future, when these men have been followed for a longer period, we expect that death due to cardiovascular diseases and cancers play a larger role and death due to current external causes play a lesser role. A final drawback is that although military conscription was mandatory in Sweden, men with severe mental disabilities or severe chronic diseases were exempt from the military examination. Thus, our results only apply to those who did not have severe mental or chronic conditions at age 18.

Our IPW model with family fixed effects controls for unobserved factors at the family level. A drawback of using family fixed effects models is that it only relies on within-brother variation in education (and other covariates in the IPW) to identify the educational gain. Brothers with the same education do not contribute to the estimation, nor do brothers who differ more than one education level. If for affluent families the main reason for a difference in education, most likely in the upper end of the distribution, between brothers is a health difference the family fixed effects model will overestimate the educational gain. Another issue is that unobserved early life health differences between brothers may still exist and influence both schooling and mortality. First, we assume a time-invariant family effect, excluding any impact of a change in family income and wealth for a later born brother. Second, although we include important individual confounders such as cognitive ability and psychological fitness, still may have excluded relevant individual characteristics that influence both schooling and mortality, e.g. low birth weight, we do not observe.

We have not investigated the mechanisms that explain the educational difference in cause-specific mortality. Ross and Wu (1995) have shown that education shapes work and economic conditions (labor market participation, poverty and work fulfillment), social-psychological resources (sense of control and social support) and healthy lifestyles (smoking, exercise, drinking and health check-ups). Healthy lifestyles are strongly related to parental SES and education and are, therefore, largely accounted for. Education also increases the individual skills to acquire knowledge and information to understand the doctor better, obtain better treatment and translate this into a healthy behavior.

References

- Aalen, O. and S. Johansen (1978). An empirical transition matrix for nonhomogeneous markov chains based on censored observations. *Scandinavian Journal of Statistics* 5, 141–150.
- Albouy, V. and L. Lequien (2009). Does compulsory education lower mortality? *Journal of Health Economics* 28(1), 155–168.

- Amin, V., J. R. Behrman, and H. P. Kohler (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.
- Andersen, P. K. (2013). Decomposition of number of life years lost according to causes of death. *Statistics in Medicine* 32, 5278–5285.
- Andersen, P. K. and V. Canudas-Romo (2013). Cause-specific measures of life years lost. *Demographic Research* 29, 1127–1152.
- Arendt, J. N. (2005). Does education cause better health? a panel data analysis using school reforms for identification. *Economics of Education Review* 24(2), 149–160.
- Austin, P. C. (2014). A tutorial on the use of propensity score methods with survival or time-to-event outcomes: Reporting measures of effect similar to those used in randomized experiments. *Statistics in Medicine* 33, 1242–1258.
- Backlund, E., P. D. Sorlie, and N. J. Johnson (1999). A comparison of the relationships of education and income with mortality: The National Longitudinal Mortality Study. *Social Science & Medicine* 49, 1373–1384.
- Behrman, J., H. Kohler, V. Jensen, D. Pedersen, I. Petersen, P. Bingley, and K. Christensen (2011). Does more schooling reduce hospitalization and delay mortality? new evidence based on danish twins. *Demography* 48(4), 1347–1375.
- Beyersmann, J. and T. H. Scheike (2013). Classical regression models for competing risks. In J. P. Klein, H. C. Van Houwelingen, J. G. Ibrahim, and T. H. Scheike (Eds.), *Handbook of Survival Analysis*, Chapter 8, pp. 157–177. Boca Raton: CRC Press.
- Bijwaard, G. E. and A. M. Jones (2016). Cognitive ability and the mortality gradient by education: Selection or mediation? Discussion Paper 9798, IZA.
- Bijwaard, G. E., M. Myrskylä, P. Tynelius, and F. Rasmussen (2016). Education, cognitive ability and cause-specific mortality: A structural approach. Working Paper 2016-007, MPIDR.
- Bijwaard, G. E., H. van Kippersluis, and J. Veenman (2015). Education and health: the role of cognitive ability. *Journal of Health Economics* 42, 29–43.
- Bijwaard, G. E., F. van Poppel, P. Ekamper, and L. H. Lumey (2015). Gains in life expectancy associated with higher education in men. *PLoS ONE* 10, e0141200.
- Borrell, C., A. Plasència, M. Huisman, G. Costa, A. Kunst, O. Andersen, M. Bopp, J.-K. Borgan, P. Deboosere, M. Glickman, S. Gadeyne, C. Minder, E. Regidor, T. Spadea, T. Valkonen, and J. P. Mackenbach (2005). Education level inequalities and transportation injury mortality in the middle aged and elderly in european settings. *Injury Prevention* 11(3), 138–142.
- Caliendo, M. and S. Kopeinig (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys* 22, 31–72.
- Ceci, S. J. (1991). How much does schooling influence general intelligence and its cognitive components? a reassessment of the evidence. *Developmental Psychology* 27, 703–722.
- Clark, D. and H. Royer (2013). The effect of education on adult mortality and health: Evidence from Britain. *American Economic Review* 103(6), 2087–2120.
- Cole, S. R. and M. A. Hernán (2004). Adjusted survival curves with inverse probability weights. *Computer Methods and Programs in Biomedicine* 75, 45–49.
- Conti, G., J. J. Heckman, and S. Urzua (2010). The education-health gradient. *American Economic Review* 100, 234–238.
- Cox, D. R. (1972). Regression models and life-tables (with discussion). *Journal of the Royal Statistical Society: Series B* 34, 187–220.

- Cutler, D. M., A. Lleras-Muney, and T. Vogl (2011). Socioeconomic status and health: Dimensions and mechanisms. In S. Glied and P. C. Smith (Eds.), *The Oxford Handbook of Health Economics*, pp. 124–163. Oxford: Oxford UP.
- Efron, B. (1982). *The jackknife, the bootstrap, and other resampling plans*. Philadelphia, PA: Society for Industrial and Applied Mathematics.
- Elo, I. T., P. Martikainen, and M. Myrskylä (2014). Socioeconomic status across the life course and all-cause and cause-specific mortality in Finland. *Social Science & Medicine* 119, 198–206.
- Elo, I. T. and S. H. Preston (1996). Educational differentials in mortality: United States, 1979–85. *Social Science & Medicine* 42, 47–57.
- Feng, P., X.-H. Zhou, Q.-M. Zou, M.-Y. Fan, and Z.-S. Li (2012). Generalized propensity score for estimating the average treatment effect of multiple treatments. *Statistics in Medicine* 31, 681–697.
- Fine, J. P. and R. J. Gray (1999). A proportional hazards model for the subdistribution of a competing risk. *Journal of the American Statistical Association* 94, 496–509.
- 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.
- Galobardes, B., J. W. Lynch, and G. D. Smith (2004). Childhood socioeconomic circumstances and cause-specific mortality in adulthood: systematic review and interpretation. *Epidemiologic Reviews* 26(1), 7–21.
- Geyer, S., Ö. Hemström, R. R. Peter, and D. Vågerö (2006). Education, income, and occupational class cannot be used interchangeably in social epidemiology. empirical evidence against a common practice. *Journal of epidemiology and community health* 60(9), 804–810.
- Grossman, M. (2015). The relationship between health and schooling: What’s new? *Nordic Journal of Health Economics* 3(1), 7–17.
- Hirano, K., G. W. Imbens, and G. Ridder (2003). Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica* 71, 1161–1189.
- Huisman, M., A. E. Kunst, M. Bopp, J.-K. Borgan, C. Borrell, G. Costa, P. Deboosere, S. Gadeyne, M. Glickman, C. Marinacci, C. Minder, T. Valkonen, and J. P. Mackenbach (2005). Educational inequalities in cause-specific mortality in middle-aged and older men and women in eight western European populations. *The Lancet* 365(9458), 493–500.
- Hummer, R. A. and J. T. Lariscy (2011). Educational attainment and adult mortality. In R. G. Rogers and E. M. Grimmins (Eds.), *International Handbook of Adult Mortality*, Volume 2, pp. 241–260. Dordrecht, the Netherlands: Springer.
- Hummer, R. A., R. G. Rogers, and I. W. Eberstein (1998). Sociodemographic differentials in adult mortality: A review of analytic approaches. *Population and Development Review* 24(3), 553–578.
- Imai, K. and D. A. van Dyk (2004). Causal inference with general treatment regimes: Generalizing the propensity score. *Journal of the American Statistical Association* 99, 854–866.
- Imbens, G. W. (2004). Nonparametric estimation of average treatment effects under exogeneity. *The Review of Economics and Statistics* 86, 4–29.
- Kulhánová, I., R. Hoffmann, T. A. Eikemo, G. Menielle, and J. P. J. P. Mackenbach (2014). Educational inequalities in mortality by cause of death: first national data for the Netherlands. *International journal of public health* 59(5), 687–696.
- Lechner, M. (2002). Some practical issues in the evaluation of heterogeneous labour market programmes by matching methods. *Journal of the Royal Statistical Society: Series A* 165, 59–82.

- Leuven, E., E. Plug, and M. Rønning (2016). Education and cancer risk. *Labour Economics* 43, 106–121.
- Link, B. G. and J. C. Phelan (1995). Social conditions as fundamental causes of disease. *Journal of Health and Social Behavior Extra issue*, 80–94.
- Lleras-Muney, A. (2005). The relationship between education and adult mortality in the United States. *Review of Economic Studies* 72, 189–221.
- Lorant, V., A. E. Kunst, M. Huisman, G. Costa, and J. P. Mackenbach (2005). Socio-economic inequalities in suicide: a European comparative study. *The British journal of psychiatry* 187(1), 49–54.
- Lundborg, P. (2013). The health returns to schooling what can we learn from twins? *Journal of Population Economics* 26(2), 673–701.
- Lundborg, P., C. H. Lyttkens, and P. Nystedt (2016). The effect of schooling on mortality: New evidence from 50,000 Swedish twins. *Demography* 53(4), 1135–1168.
- Lynch, S. M. (2003). Cohort and life-course patterns in the relationship between education and health: A hierarchical approach. *Demography* 40(2), 309–333.
- Mackenbach, J. P. (2012). The persistence of health inequalities in modern welfare states: The explanation of the paradox. *Social Science & Medicine* 75, 761–769.
- Mackenbach, J. P., I. Kulháňová, M. Bopp, P. Deboosere, T. A. Eikemo, R. Hoffmann, M. C. Kulik, M. Leinsalu, P. Martikainen, G. Menvielle, E. Regidor, B. Wojtyniak, O. Östergren, and O. Lundberg (2015). Variations in the relation between education and cause-specific mortality in 19 European populations: A test of the fundamental causes theory of social inequalities in health. *Social Science & Medicine* 127, 51–62.
- Mackenbach, J. P., I. Stirbu, A.-J. R. Roskam, M. M. Schaap, G. Menvielle, M. Leinsalu, and A. E. Kunst (2008). Socioeconomic inequalities in health in 22 European countries. *New England Journal of Medicine* 358(23), 2468–2481.
- Masters, R. K., R. A. Hummer, and D. A. Powers (2012). Educational differences in US adults mortality: A cohort perspective. *American Sociological Review* 77(4), 548–572.
- Mazumder, B. (2008). Does education improve health: A reexamination of the evidence from compulsory schooling laws. *Economic Perspectives* 33, 2–16.
- Meghir, C., M. Palme, and E. Simeonova (2013). Education, cognition and health: Evidence from a social experiment. Working paper no. 19002, NBER.
- Mirowsky, J. and C. E. Ross (2003). *Education, Social Status, and Health*. New York: Aldine de Gruyter.
- Montez, J. K., R. Hummer, and M. D. Hayward (2012). Educational attainment and adult mortality in the United States: A systematic analysis of functional form. *Demography* 49, 315–336.
- Næss, Ø., D. A. Hoff, D. Lawlor, and L. H. Mortensen (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.
- Overgaard, M., P. K. Andersen, and E. Parner (2015). Regression analysis of censored data using pseudo-observations: An update. *Stata Journal* 15, 809–821.
- Parner, E. and P. K. Andersen (2010). Regression analysis of censored data using pseudo-observations. *Stata Journal* 10, 408–422.
- Preston, S. H. and P. Taubman (2011). Socioeconomic differences in adult mortality and health status. In L. G. Martin and S. H. Preston (Eds.), *Demography of aging*, pp. 279–318. Washington, DC: National Academy Press.

- Putter, H., M. Fiocco, and R. B. Geskus (2007). Tutorial in biostatistics: Competing risks and multi-state models. *Statistics in Medicine* 26, 2389–2430.
- Rosenbaum, P. and D. B. Rubin (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika* 70, 41–50.
- Ross, C. E. and C.-L. Wu (1995). The links between education and health. *American Sociological Review* 60(5), 719–745.
- Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and non-randomized studies. *Journal of Educational Psychology* 66, 688–701.
- Smith, J. P. (2004). Unraveling the SES-health connection. *Population and Development Review* 30, 108–132.
- Thomas, J. M. (1996). On the interpretation of covariate estimates in independent competing risks models. *Bulletin of Economic Research* 48, 27–39.
- Van Kippersluis, H., O. O'Donnell, and E. van Doorslaer (2011). Long run returns to education: Does schooling lead to an extended old age? *Journal of Human Resources* 46(4), 695–721.
- Xiang, F. and S. Murray (2012). Restricted mean models for transplant benefit and urgency. *Statistics in Medicine* 31, 561–576.

Appendix A Methodology

The standard approach to analyse cause-specific mortality is to formulate a competing risks model with an independent Cox proportional hazard model for each cause-specific mortality. The cause-specific hazard of dying from cause j for an individual with characteristics X is

$$\lambda_j(t|X) = \lambda_{0j}(t) \exp(\beta_j' X),$$

with $\lambda_{0j}(t)$ is the age-dependence of the mortality rate, common for all individuals. The cause-specific hazard gives the mortality rate due a particular cause conditional on not having died from any other cause previously. The interpretation of the coefficients in such a competing risks model requires caution, Thomas (1996). A particular covariate, say x_l , can appear in several competing hazards. In such a case the vectors β_{lj} convey little information about the effect of the covariate on the probability to die from cause j , because that probability depends not only on the hazard to die from cause j but also on the hazard to die from all the other causes. Thus, despite that in a Cox analysis for competing causes of death it is assumed that the causes are independent many measures of the importance of the specific causes depend on all other hazards.

Note that total hazard of dying is the sum of all cause-specific hazards and the total survival function, the probability to survive up to age t , is equal to

$$S(t|X) = \exp\left(-\sum_j \int_0^t \lambda_j(s|X) ds\right). \quad (\text{A.1})$$

Another approach to model competing risks is to estimate a regression model for the cumulative incidence function,

$$F_j(t|X) = \int_0^t \lambda_j(s|X) S(s|X) ds, \quad (\text{A.2})$$

which is the probability of dying from cause j before time t . Note that the cumulative probability to die from cause j remains below one, as the sum all cumulative incidence functions is equal to one minus the survival function. The Fine and Gray (1999) regression model for the cumulative incidence assumes that the cumulative incidence for cause j is given by

$$F_j(t|X) = 1 - \exp\left(-\Lambda_0(t)e^{\beta_j' X}\right) \quad (\text{A.3})$$

where β_j is the vector of regression coefficients related to cause j and $\Lambda_0(t)$ is a unspecified non-decreasing function with $\Lambda_0(0) = 0$. This model has some resemblance with the Cox proportional hazard model. However, the interpretation of the coefficients is rather difficult. First, the Fine–Gray model regresses the covariates on $\ln(-\ln(1 - F_j(t|X)))$. A more substantive issue with the Fine–Gray model is that it is based on the ‘subdistribution hazard’, which is not the cause-specific hazard. Implicit in the subdistribution hazard is that it assumes that any individual who dies from a competing cause remains in the risk set of a specific cause. This implies that the risk sets associated with the subdistribution hazard will be unknown for competing events after they have occurred.⁷

A direct way to measure the impact of education on cause-specific mortality is to calculate the months lost due to a specific cause. The months lost due to cause j is (from τ_0 to τ_1 , e.g. from age 18 till age 63) are directly related to the cumulative incidence functions

$$L_j(\tau_0, \tau_1) = \int_{\tau_0}^{\tau_1} F_j(s) ds \quad (\text{A.4})$$

This measure is related to the mean life time, or the restricted mean life time (from τ_0 to τ_1), see Andersen (2013). The sum over all causes of the months-lost due to each specific cause is equal to the

⁷For a detailed discussion on these issues see Putter et al. (2007) or Beyersmann and Scheike (2013).

total expected months lost between τ_0 and τ_1 , which is τ_1 minus the restricted mean life time. Non-parametric estimation of the months-lost is straightforward because the survival (Kaplan-Meier) and cumulative incidence functions (Aalen-Johansen, Aalen and Johansen 1978) are also straightforward to calculate. If we do not have ties in observed ages of death the Kaplan-Meier estimator of the, total, survival is

$$S(t) = \prod_{s \leq t} \left(1 - \frac{d(s)}{Y(s)}\right) \quad (\text{A.5})$$

with $d(s)$ is the number of deaths at age s and $Y(s)$ is the number of people still alive at age s . From the Nelson-Aalen estimate of the cumulative hazard for cause j ,

$$\Lambda_j(t) = \sum_{s \leq t} \frac{d_j(s)}{Y(s)} \quad (\text{A.6})$$

with $d_j(s)$ is the number of deaths due to cause j at age s , the Aalen-Johansen estimate of the cumulative incidence of cause j is

$$F_j(t) = \sum_{s \leq t} S(s_-) [\Lambda_j(s) - \Lambda_j(s_-)] \quad (\text{A.7})$$

with s_- is the age just before s and an estimate of the months lost is

$$L_j(\tau_0, \tau_1) = \sum_{\tau_0 \leq s \leq \tau_1} F_j(s_-)(s - s_-) \quad (\text{A.8})$$

Appendix A.1 Using pseudo observations

If all individuals were observed till they the impact of education on the month-lost would be very easy to obtain. However, common for survival analysis, the individual deaths are (heavily) censored at the end of the observation window 1/1/2013. This implies that the highest age reached we can observe is 63 years. We will use pseudo-observations to account for censoring. Andersen (2013) has shown that with pseudo-observations regression analysis for the months-lost due to a specific cause of death is very simple, even when some of the observations are censored. The idea of pseudo-observations is closely linked to the Jackknife-method, (Efron 1982). If the sample contains n individuals the jackknife uses n deterministically defined subsamples of size $n - 1$ obtained by dropping in turn each of the n observations and re-estimate the model. The advantage of creating pseudo-observations is that they can be modelled using standard (uncensored) linear models. For the estimation of the years lost the pseudo-observation for each individual is calculated as

$$J_k^i = n \hat{L}_k(18, 63) - (n - 1) \hat{L}_k^{(-i)}(18, 65), \quad i = 1, \dots, n \quad (\text{A.9})$$

where $\hat{L}_k^{(-i)}(18, 63)$ is the estimator of months-lost from the sample without the i^{th} individual and $\hat{L}_k(18, 65)$ the full sample estimate. We regress these obtained months-lost on the education indicator using a GLM approach.⁸

Appendix A.2 Inverse propensity score weighting

We define the treatment effect, of moving up one education level, in terms of a proportional change in the (mortality) hazard rate. 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

⁸Parner and Andersen (2010) and Overgaard et al. (2015) provide an STATA procedure to estimate the years/months lost based on pseudo observations.

risk. The main difference with standard propensity score methods is that we use potential hazard rates, the hazard rate that would be observed if the individual was untreated, $\lambda(t|0)$, or treated $\lambda(t|1)$. Let $D_i = 1$ be the treatment, moving up one education level. We observe pre-treatment (educational level) covariates X that influence the education choice.

Assumption 1. *Unconfoundedness:* $\lambda(t|d) \perp D|X$ for $d = 0, 1$

where \perp denotes independence. The unconfoundedness assumption (Rubin 1974, Rosenbaum and Rubin 1983) asserts that, conditional on covariates X , treatment assignment (education level) is independent of the potential outcomes. This assumption requires that all variables that affect both the mortality and the education choice are observed. Note that this does not imply that we assume all relevant covariates are observed. Any missing factor is allowed to influence either the outcome or the education choice, not both. Although this is not testable and clearly a strong assumption, it may be a reasonable approximation. Any alternative, that does not rely on unconfoundedness while allowing for consistent estimation of the average treatment effects, will have to make alternative untestable assumptions.

Assumption 2. *Overlap:* $0 < \Pr(D = 1|X) < 1$.

The overlap, or common support assumption requires that the propensity score, the conditional probability to choose a higher education given covariates X is bounded away from zero and one. This assumption is in principle testable. If there are values of the covariates for which the probability of choosing a higher education level is zero or one, we cannot compare the ‘treated’ and ‘control’ individuals at these values. In that case we have to limit comparisons to sets of values where there is sufficient mass in the propensity score among both treated and controls. By comparing only adjacent education levels we remove the overlap problems.

Rosenbaum and Rubin (1983) show that if the potential outcomes are independent of treatment conditional on covariates X , they are also independent of treatment conditional on the propensity score, $p(x) = \Pr(D = 1|X = x)$. 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. Here we use weighting on the propensity score. Inverse probability weighting based on the propensity score creates a pseudo-population in which the educational attainment is independent of the measured confounders. The pseudo-population is the result of assigning to each individual a weight that is proportional to the inverse of their propensity score. Inverse probability weighting (IPW) estimation is usually based on normalized weights that add to unity. Suppose we have a sample of n individuals, then based on an estimate of the propensity score, $\hat{p}(x)$, the estimator of the average treatment effect is of the form $\sum_i [W_i \cdot D_i \cdot Y_i - W_i \cdot (1 - D_i) \cdot Y_i]$ with weights normalised to one:

$$W_i = \left[\frac{D_i}{\hat{p}(X_i)} \right] / \left[\sum_{j=1}^n \frac{D_j}{\hat{p}(X_j)} \right] + \left[\frac{(1 - D_i)}{1 - \hat{p}(X_i)} \right] / \left[\sum_{j=1}^n \frac{1 - D_j}{1 - \hat{p}(X_j)} \right] \quad (\text{A.10})$$

We implement an IPW-version of the estimation procedure for months-lost based on pseudo-observations. Xiang and Murray (2012) developed a related method for estimation of the restricted mean. The estimation of the IPW-months lost is straightforward and based on an extension of the Stata procedure `stplost` of Overgaard et al. (2015). After calculating the weights, based on the estimated propensity scores, we construct a pseudo population. Then for this pseudo population we estimate the total survival-function, using a Kaplan-Meier estimate, and the cumulative incidence function, using the Aalen-Johansen method, and use the latter in (A.4) to estimate the base months-lost for each cause of death. Next we construct the pseudo-observations, using (A.9) and regress these

IPW-pseudo observations on the education indicator to obtain our estimate of the impact of education on months-lost.

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 both the control and treated group. One suitable way to check whether there are still differences is by calculating the standardized bias, or normalised difference in means:

$$100 \cdot \frac{\bar{x}_1 - \bar{x}_0}{\sqrt{0.5(\text{Var}(x_1) + \text{Var}(x_0))}} \quad (\text{A.11})$$

Table B.5 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.

We have also checked whether the propensity score have overlap for each adjacent education level, see Figure B.3 in Appendix B. We do not observe common support problems.

Appendix B Additional tables and figure

Table B.1: Cox hazard estimates education impact cause-specific mortality (HR), full sample: men with and without brothers

	Secondary education (2 years)	Post-secondary (3 years)	Post-secondary (< 3 years)	Higher
<i>Cancers</i>				
bowel cancer	0.904	0.791 ⁺	1.081	1.039
digestive cancer	0.826**	0.840 ⁺	0.790 ⁺	0.687**
lung cancer	0.718**	0.737**	0.770	0.707 ⁺
leukemia	0.636**	0.932	1.160	0.642 ⁺
lymphoma	1.034	0.925	1.044	0.898
other cancer	0.851**	0.926	0.962	0.789**
<i>CVD</i>				
Ischemic heart disease	0.675**	0.756**	0.705**	0.672**
stroke	0.812**	0.628**	0.881	0.732 ⁺
other CVD	0.789**	0.763**	0.762**	0.569**
<i>Other diseases</i>				
Infectious	0.664**	0.424**	1.019	0.841
endocrine	0.598**	0.786	0.689	0.869
diabetes	0.686**	0.660**	0.705 ⁺	0.480**
psychiatric	0.498**	0.430**	0.388**	0.734
nervous system	0.576**	0.829	0.774	0.756
respiratory disease	0.575**	0.566**	0.530**	0.820
digestive disease	0.591**	0.530**	0.765	0.930
liver	0.674**	0.622**	0.515**	0.618**
abnormalities	0.516**	0.986	0.675**	0.845
other natural causes	0.577**	0.889	0.698**	0.877
<i>External causes</i>				
traffic accidents	0.703**	0.729**	0.583**	0.428**
suicide	0.680**	0.735**	0.779**	0.553**
other external causes	0.550**	0.649**	0.616**	0.454**

The full sample contains 1,570,567 individuals (men) of which 205,739 belongs to the lowest education group; 476,709 has finished secondary education (2 years); 326,276 has finished full secondary education (3 years); 231,102 has post-secondary education and 330,741 has attained higher education. Separate models are estimated for each two adjacent education levels. ⁺ $p < 0.05$, ** $p < 0.01$.

Table B.2: Descriptive statistics: men without (known) brothers

	Primary	Secondary education (2 years)	(3 years)	Post-secondary (< 3 years)	Higher
<i>SES mother at birth</i>					
non-manual (high)	1%	1%	2%	4%	7%
non-manual (intermediate)	3%	4%	9%	12%	18%
non-manual (low)	14%	19%	21%	28%	29%
Farmers	11%	7%	4%	5%	3%
Skilled workers	32%	33%	16%	20%	12%
Unskilled workers	20%	19%	22%	15%	12%
Not classified	17%	14%	25%	15%	18%
Unknown	1%	1%	1%	1%	1%
<i>Education mother</i>					
Primary (< 9 yrs)	49%	44%	20%	25%	15%
Primary (9–10 yrs)	13%	12%	17%	12%	10%
Secondary education (2 yrs)	25%	32%	39%	33%	29%
Secondary education (3 yrs)	3%	3%	7%	7%	8%
Post-secondary	3%	5%	9%	12%	16%
Higher	2%	2%	5%	10%	19%
Unknown	4%	2%	2%	2%	2%
<i>SES father at birth</i>					
non-manual (high)	2%	2%	4%	7%	15%
non-manual (intermediate)	8%	10%	15%	23%	27%
non-manual (low)	7%	8%	10%	12%	12%
Farmers	8%	5%	5%	4%	4%
Skilled workers	26%	29%	26%	22%	16%
Unskilled workers	36%	35%	29%	23%	16%
Not classified	8%	6%	7%	7%	7%
Unknown	5%	4%	3%	3%	3%
<i>Education father</i>					
Primary (< 9 yrs)	48%	46%	25%	26%	17%
Primary (9–10 yrs)	9%	7%	14%	7%	6%
Secondary education (2 yrs)	18%	21%	27%	21%	18%
Secondary education (3 yrs)	7%	10%	14%	18%	18%
Post-secondary	3%	4%	7%	9%	11%
Higher	2%	2%	6%	11%	24%
Unknown	13%	9%	6%	7%	6%
mother < 20 at birth	13%	9%	5%	5%	3%
father > 40 at birth	13%	11%	7%	9%	9%
birth order	2.5	2.4	2.2	2.1	2.1
global IQ ^a	3.8	4.4	4.9	6.1	6.6
Psychological assessment ^a	4.1	4.7	5.0	5.5	5.7
# of individuals	106,759	240,538	179,227	122,123	172,545

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

Table B.3: Months lost due to a specific cause 18-63 by education level, full sample unadjusted

	Primary	Educational gain			
		Secondary education (2 years)	(3 years)	Post-secondary (< 3 years)	Higher
<i>Cancers</i>	3.43**	0.76**	0.29**	0.16 ⁺	0.38**
bowel cancer	0.36**	0.05	0.04	-0.01	-0.01
digestive cancer	0.75**	0.17**	0.09 ⁺	0.10**	0.08 ⁺
lung cancer	0.51**	0.19**	0.07**	0.07**	0.02
leukemia	0.24**	0.11**	0.01	-0.03	0.06**
lymphoma	0.35**	-0.00	0.01	-0.00	0.04
other cancer	1.22**	0.25**	0.06	0.02	0.19**
<i>CVD</i>	3.95**	1.39**	0.54**	0.56**	0.29**
Ischemic heart disease	2.12**	0.88**	0.25**	0.31**	0.08 ⁺
stroke	0.51**	0.13**	0.12**	0.05	0.03
other CVD	1.32**	0.39**	0.17**	0.21**	0.18**
<i>Other diseases</i>	5.75**	2.90**	0.80**	0.85**	0.13 ⁺
Infectious	0.34**	0.15**	0.09**	0.02	0.00
endocrine	0.23**	0.10**	0.02	0.04 ⁺	0.00
diabetes	0.40**	0.15**	0.08**	0.05 ⁺	0.05 ⁺
psychiatric	1.46**	0.85**	0.29**	0.22**	-0.02
nervous system	0.64**	0.32**	0.03	0.08**	0.03
respiratory disease	0.62**	0.32**	0.10**	0.10**	-0.00
digestive disease	0.24**	0.11**	0.05**	0.03 ⁺	-0.01
liver	0.79**	0.33**	0.15**	0.16**	0.00
abnormalities	0.80**	0.46**	-0.02	0.11**	0.05
other natural causes	0.23**	0.10**	0.01	0.05**	0.03 ⁺
<i>External causes</i>	12.19**	5.35**	1.68**	2.02**	1.75**
traffic accidents	2.72**	1.13**	0.37**	0.61**	0.49**
suicide	4.31**	1.65**	0.61**	0.57**	0.71**
other external causes	5.16**	2.58**	0.71**	0.85**	0.56**
total	25.31	10.40**	3.32**	3.59**	2.55**

⁺ $p < 0.05$, ** $p < 0.01$.

Table B.4: Logit estimates of propensity scores, pairwise comparisons

	Educational levels ^a			
	(1)	(2)	(3)	(4)
<i>Mother's SES at birth</i>				
Others, not classified	-0.036 ⁺	0.053**	-0.236**	0.095**
Unskilled workers	0.010	0.030 ⁺	-0.176**	0.057**
Farmers	-0.238**	0.013	-0.088**	0.054 ⁺
Non-manual workers at lower level	0.108**	0.201**	-0.109**	0.120**
Non-manual workers at intermediate level	0.060 ⁺	0.140**	-0.089**	0.120**
Non-manual workers at higher level	-0.036	0.437**	-0.167**	0.284**
missing	0.050	0.200**	-0.043	0.155**
<i>Father's SES at birth</i>				
Others, not classified	-0.189**	-0.005	0.105**	0.066**
Unskilled workers	-0.086**	-0.028**	-0.018	0.003
Farmers	-0.021	0.042	0.005	-0.072**
Non-manual workers at lower level	-0.057**	0.272**	0.068**	0.104**
Non-manual workers at intermediate level	-0.000	0.164**	0.203**	0.109**
Non-manual workers at higher level	-0.131**	0.262**	0.246**	0.236**
missing	0.071	0.143**	0.059	0.144**
<i>Mother's education</i>				
less than 9 years	-0.208**	-0.105**	-0.047**	-0.138**
9-10 years	-0.138**	-0.008	-0.120**	-0.044**
Full secondary education	-0.060 ⁺	0.157**	0.110**	0.129**
University, less than 3 years	0.027	0.169**	0.221**	0.192**
University, at least 3 years	-0.112**	0.291**	0.300**	0.340**
PhD studies	-0.294	0.555**	0.188	0.556**
missing	-0.257**	-0.108 ⁺	-0.014	-0.045
<i>Father's education</i>				
less than 9 years	-0.194**	-0.127**	-0.034**	-0.067**
9-10 years	-0.138**	-0.046**	-0.100**	-0.053**
Full secondary education	0.050**	0.153**	0.192**	0.004
University, less than 3 years	0.087**	0.202**	0.251**	0.093**
University, at least 3 years	-0.055	0.470**	0.263**	0.460**
PhD studies	-0.256 ⁺	0.835**	0.244**	0.731**
missing	-0.169**	-0.107**	0.075 ⁺	0.071 ⁺
constant	1.851**	-1.369**	0.043	-0.645**

^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.

Reference category: mother/father skilled worker and secondary education (max 2 years), IQ level and psychological assessment 5. County and year dummies are also included. ⁺ $p < 0.05$, ** $p < 0.01$.

Table B.4: Logit estimates of propensity scores, pairwise comparisons (continued)

	Educational levels ^a			
	(1)	(2)	(3)	(4)
Mother < 20 at birth	-0.222**	-0.246**	-0.134**	-0.174**
Father > 40 at birth	0.106**	0.120**	0.078**	0.117**
missing father	-0.037	-0.178	-0.048	-0.261 [†]
birth order 2	-0.096**	-0.132**	-0.066**	0.042**
birth order 3	-0.177**	-0.184**	-0.097**	0.024
birth order 4	-0.237**	-0.239**	-0.139**	0.035
birth order ≥ 5	-0.324**	-0.344**	-0.098**	-0.037
Global IQ 1 ^b	-0.886**	-0.905**	-0.974**	-0.209**
2	-0.633**	-0.647**	-0.858**	-0.452**
3	-0.431**	-0.494**	-0.589**	-0.303**
4	-0.219**	-0.287**	-0.296**	-0.208**
6	0.144**	0.343**	0.325**	0.138**
7	0.246**	0.674**	0.605**	0.322**
8	0.177**	1.083**	0.841**	0.524**
9	-0.063	1.643**	1.060**	0.840**
missing	-0.840**	-0.418**	0.121**	0.050
Psychological assessment 1 ^b	-0.567**	-0.542**	0.058	0.132**
2	-0.456**	-0.351**	-0.030	0.032
3	-0.321**	-0.185**	-0.137**	0.060**
4	-0.166**	-0.096**	-0.088**	-0.000
6	0.122**	0.160**	0.093**	0.073**
7	0.173**	0.336**	0.171**	0.159**
8	0.163**	0.455**	0.217**	0.238**
9	0.116 [†]	0.571**	0.210**	0.308**
missing	-0.280**	-0.091**	-0.071**	0.116**

^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.

^b Running from low to high.

Reference category: mother/father skilled worker and secondary education (max 2 years), IQ level 5 and psychological assessment level 5. County and year dummies are also included. [†] $p < 0.05$, ** $p < 0.01$.

Table B.5: Standardized bias before and after matching, pairwise comparisons

	Educational levels ^a							
	(1)		(2)		(3)		(4)	
	Before	After	Before	After	Before	After	Before	After
<i>Mother's SES at birth</i>								
Others, not classified	9.426	0.199	-40.868	0.466	32.833	1.381	-10.831	0.160
Unskilled workers	-2.307	-0.794	-5.914	-0.177	15.868	0.611	8.451	0.058
Farmers	9.619	-0.075	17.109	-0.502	-4.861	-0.441	10.206	0.067
Non-manual workers at lower level	-13.454	0.919	-5.592	-0.158	-20.449	-0.556	0.284	-0.314
Non-manual workers at intermediate level	-5.903	0.401	-19.860	0.996	-11.137	-0.314	-17.183	0.095
Non-manual workers at higher level	-0.941	0.499	-8.850	1.079	-9.453	-0.083	-17.111	0.082
missing	0.935	0.198	2.151	0.165	-1.444	-0.124	0.729	-0.031
<i>Father's SES at birth</i>								
Others, not classified	4.824	0.196	-2.608	0.964	-0.660	0.037	-3.844	-0.080
Unskilled workers	6.189	-0.492	19.735	-1.036	14.560	0.551	18.058	0.030
Farmers	7.764	-0.363	-0.036	-0.628	4.820	0.056	5.956	0.024
Non-manual workers at lower level	-3.512	0.443	-10.499	0.053	-3.205	-0.697	1.230	0.182
Non-manual workers at intermediate level	-10.637	0.950	-17.753	0.780	-19.389	-0.615	-10.933	-0.237
Non-manual workers at higher level	0.066	0.836	-16.711	1.709	-13.466	0.008	-27.630	0.162
missing	1.665	0.441	3.109	0.166	-1.422	-0.182	0.648	-0.016
<i>Mother's education</i>								
less than 9 years	16.147	-0.324	50.002	-1.474	-3.286	-0.426	26.813	-0.003
9-10 years	1.870	0.138	-13.852	0.114	13.589	1.217	7.755	0.030
Full secondary education	-3.679	0.139	-15.629	0.613	-1.280	-0.470	-5.567	-0.014
University, less than 3 years	-7.875	0.515	-19.098	0.931	-10.188	-0.518	-13.445	-0.230
University, at least 3 years	-4.041	0.612	-19.293	1.319	-17.114	0.266	-30.253	0.144
PhD studies	0.576	-0.007	-2.896	0.355	-2.470	0.117	-7.056	0.165
missing	9.099	0.241	0.704	0.019	1.913	-0.276	1.171	0.063
<i>Father's education</i>								
less than 9 years	11.795	-0.803	46.141	-1.499	5.459	0.261	26.705	-0.018
9-10 years	3.551	-0.145	-17.647	-0.209	17.645	0.869	4.556	0.034
Full secondary education (12-13)	-11.709	0.450	-14.803	0.295	-10.276	-0.723	3.278	-0.184
University, less than 3 years	-7.252	0.287	-15.641	0.611	-8.295	-0.746	-4.243	-0.054
University, at least 3 years	-3.274	1.080	-21.780	1.961	-18.495	0.043	-34.161	0.013
PhD studies	0.253	0.408	-7.303	0.603	-7.377	0.372	-16.676	0.232
missing	10.782	0.512	7.745	-0.063	-1.151	-0.371	2.301	0.006

^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.5: Standardized bias before and after matching, pairwise comparisons (continued)

	Educational levels ^a							
	(1)		(2)		(3)		(4)	
	Before	After	Before	After	Before	After	Before	After
Mother < 20 at birth	2.294	0.238	14.705	-0.311	2.311	0.047	11.017	-0.028
Father > 40 at birth	2.725	-0.005	7.400	-0.015	-1.879	-0.147	0.868	0.007
missing father	1.365	0.256	2.670	0.010	-0.548	-0.086	1.274	-0.003
birth order 2	-3.224	-0.178	-10.036	0.078	6.527	-0.083	-3.590	0.094
birth order 3	3.653	-0.077	-0.604	0.297	7.165	0.171	-0.431	0.231
birth order 4	5.668	0.104	7.753	-0.217	4.210	0.397	2.112	0.088
birth order ≥ 5	9.747	0.035	14.511	-0.304	2.272	0.184	5.250	-0.012
Global IQ 1 ^b	20.673	-0.155	10.681	-0.015	15.396	0.728	2.059	0.022
2	15.571	0.510	13.526	0.297	23.565	1.444	7.268	-0.026
3	6.296	0.292	11.577	-0.686	26.917	1.204	10.155	-0.066
4	-4.999	-0.349	6.878	-1.384	23.799	0.242	13.044	0.065
6	-16.461	-0.057	-10.515	0.278	-12.095	-1.174	5.692	0.254
7	-12.500	0.742	-15.450	1.302	-25.300	-0.718	-4.798	0.059
8	-6.031	0.634	-15.829	0.897	-27.212	-0.032	-13.642	-0.293
9	-1.278	1.021	-13.255	1.120	-22.069	0.666	-21.573	-0.685
missing	18.779	-0.094	1.070	-0.113	2.729	0.088	0.337	0.303
Psychological assessment 1 ^b	13.985	0.394	7.483	-0.277	4.603	-0.318	0.911	0.432
2	17.023	1.020	5.586	0.170	11.220	-0.132	1.103	0.511
3	9.571	0.953	8.216	0.056	12.906	0.425	2.233	0.232
4	-2.816	0.233	8.219	0.203	8.676	0.336	5.560	0.163
6	-16.821	-1.018	-4.565	-0.290	-9.024	-0.471	1.839	-0.157
7	-13.008	-0.322	-11.907	0.336	-15.022	-0.458	-6.162	-0.159
8	-7.247	0.305	-9.166	0.276	-14.344	-0.041	-8.307	-0.348
9	-3.189	0.429	-5.909	0.285	-9.184	0.298	-6.085	-0.302
missing	19.148	-0.171	-7.700	0.157	13.107	0.709	-1.525	0.428

^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.

^b Running from low to high.

Table B.6: Months lost due to a specific cause 18-63 by education level, men with brothers, IPW

	Primary	Educational gain			
		Secondary education (2 years)	(3 years)	Post-secondary (< 3 years)	Higher
<i>Cancers</i>	3.16**	0.51**	0.12	0.32 ⁺	0.41 ⁺
bowel cancer	0.33**	0.01	0.05	0.01	-0.03
digestive cancer	0.64**	0.07	0.00	0.20**	0.10
lung cancer	0.39**	0.11 ⁺	-0.00	0.08	0.09 ⁺
leukemia	0.31**	0.21**	-0.01	-0.02	0.06 ⁺
lymphoma	0.33**	-0.04	0.04	-0.06	0.08
other cancer	1.16**	0.16	0.04	0.10	0.11
<i>CVD</i>	3.32**	1.01**	0.31	0.51**	0.51**
Ischemic heart disease	1.70**	0.57**	0.16	0.30**	0.10
stroke	0.41**	0.04	0.10 ⁺	0.07	0.08 ⁺
other CVD	1.21**	0.40**	0.05	0.14	0.33**
<i>Other diseases</i>	4.69**	2.26**	0.62**	0.57**	0.27
Infectious	0.31**	0.13	0.09**	0.00	0.01
endocrine	0.15**	0.07	0.02	-0.04	0.05
diabetes	0.36**	0.15**	0.05	0.02	0.04
psychiatric	1.28**	0.76**	0.18**	0.23**	0.03
nervous system	0.58**	0.30**	0.00	0.06	0.07
respiratory disease	0.58**	0.32**	0.13**	0.06 ⁺	-0.05 ⁺
digestive disease	0.18**	0.09**	0.02	0.02	0.02
liver	0.57**	0.15**	0.15**	0.12 ⁺	0.03
abnormalities	0.54**	0.24**	0.02	0.06	0.03
other natural causes	0.17**	0.06	-0.02	0.05	0.05
<i>External causes</i>	11.96**	5.21**	1.72**	1.51**	1.55**
traffic accidents	2.71**	0.97**	0.43**	0.47**	0.39**
suicide	4.46**	1.85**	0.62**	0.50**	0.58**
other external causes	4.80**	2.39**	0.67**	0.54**	0.58**
total	23.13	8.99**	2.72**	2.91**	2.74**

⁺ $p < 0.05$, ** $p < 0.01$.

Figure B.1: Cumulative incidence cancers and cardiovascular diseases by education level

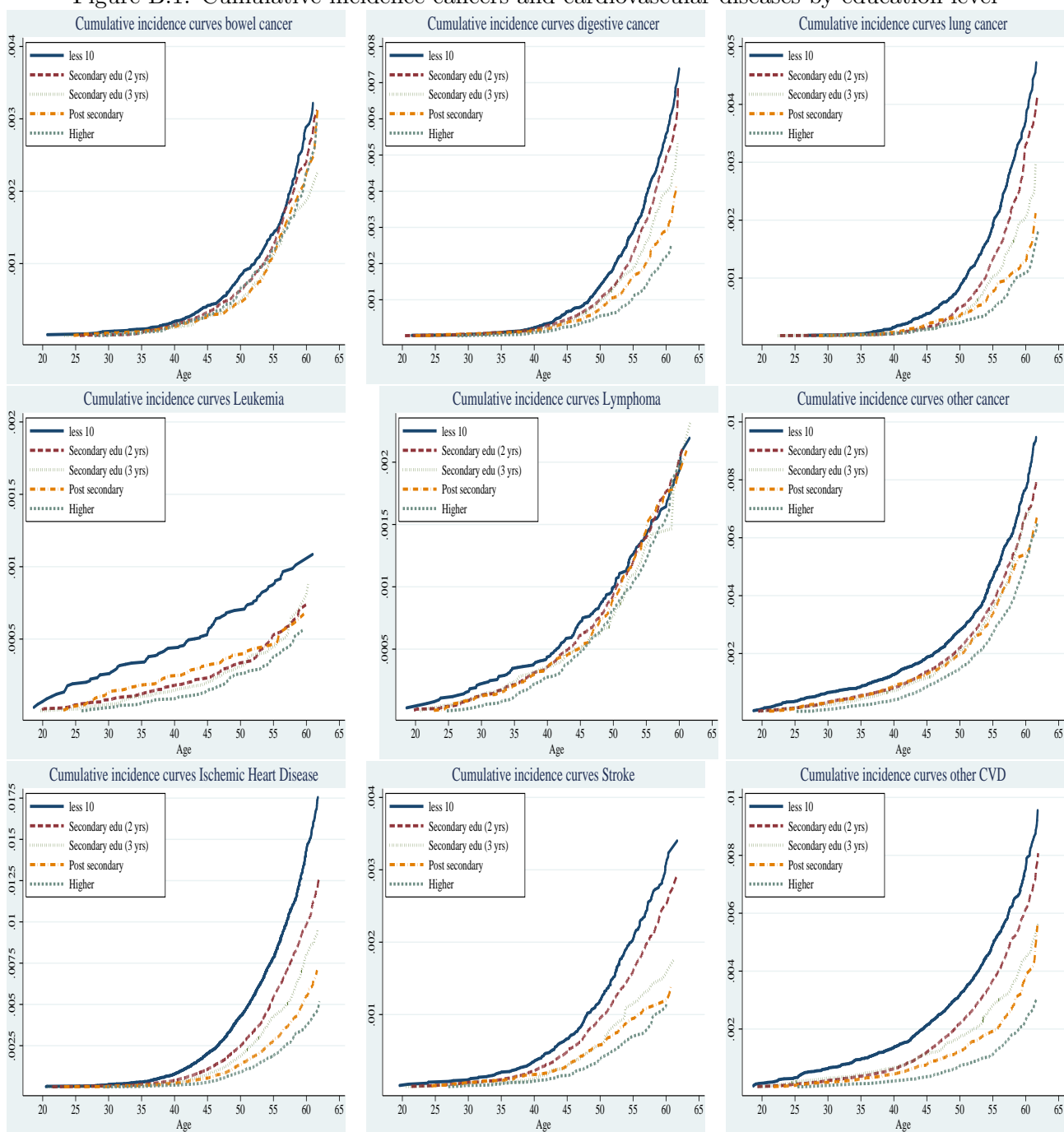


Figure B.2: Cumulative incidence other diseases by education level

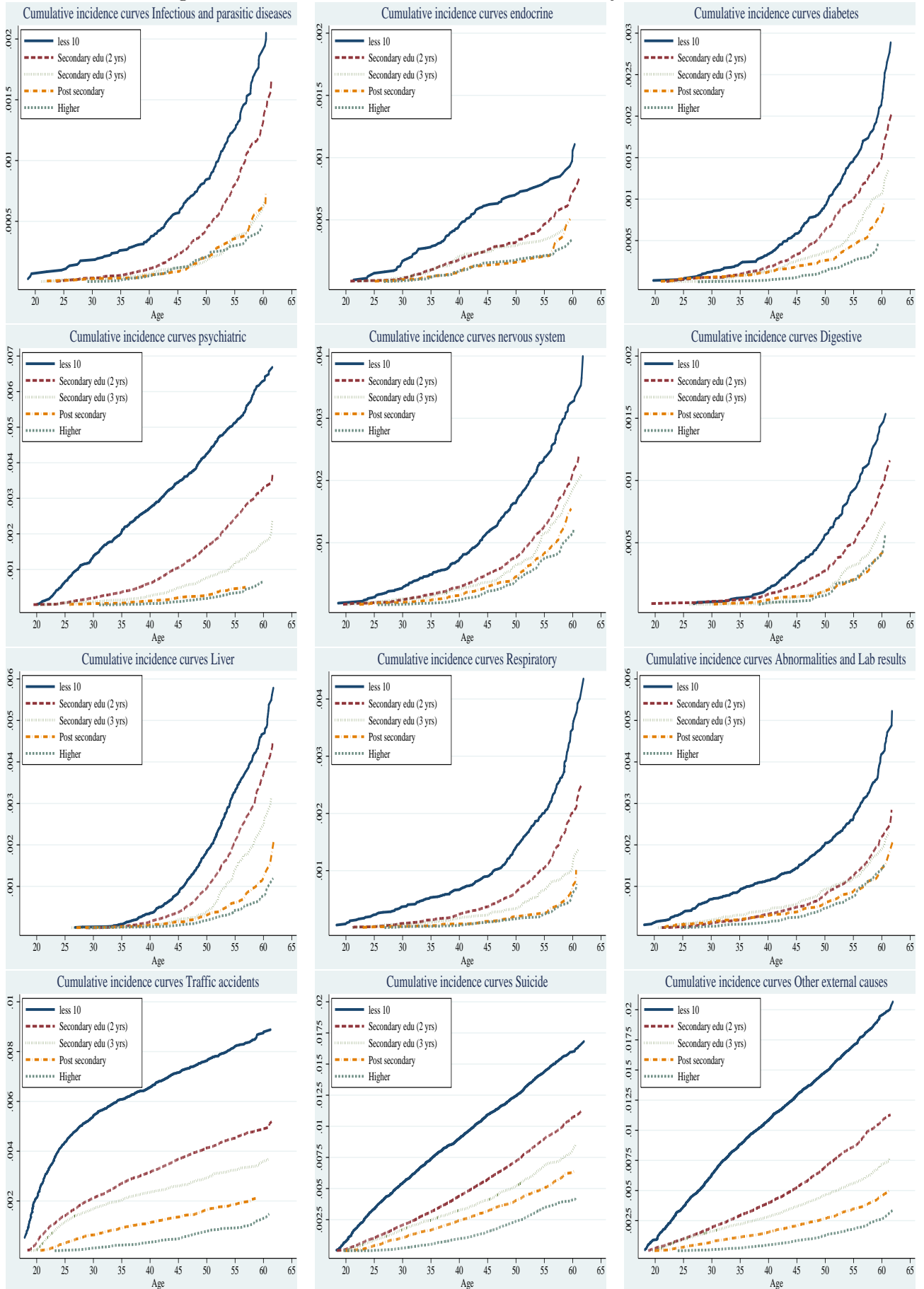


Figure B.3: Common overlap of the propensity scores

